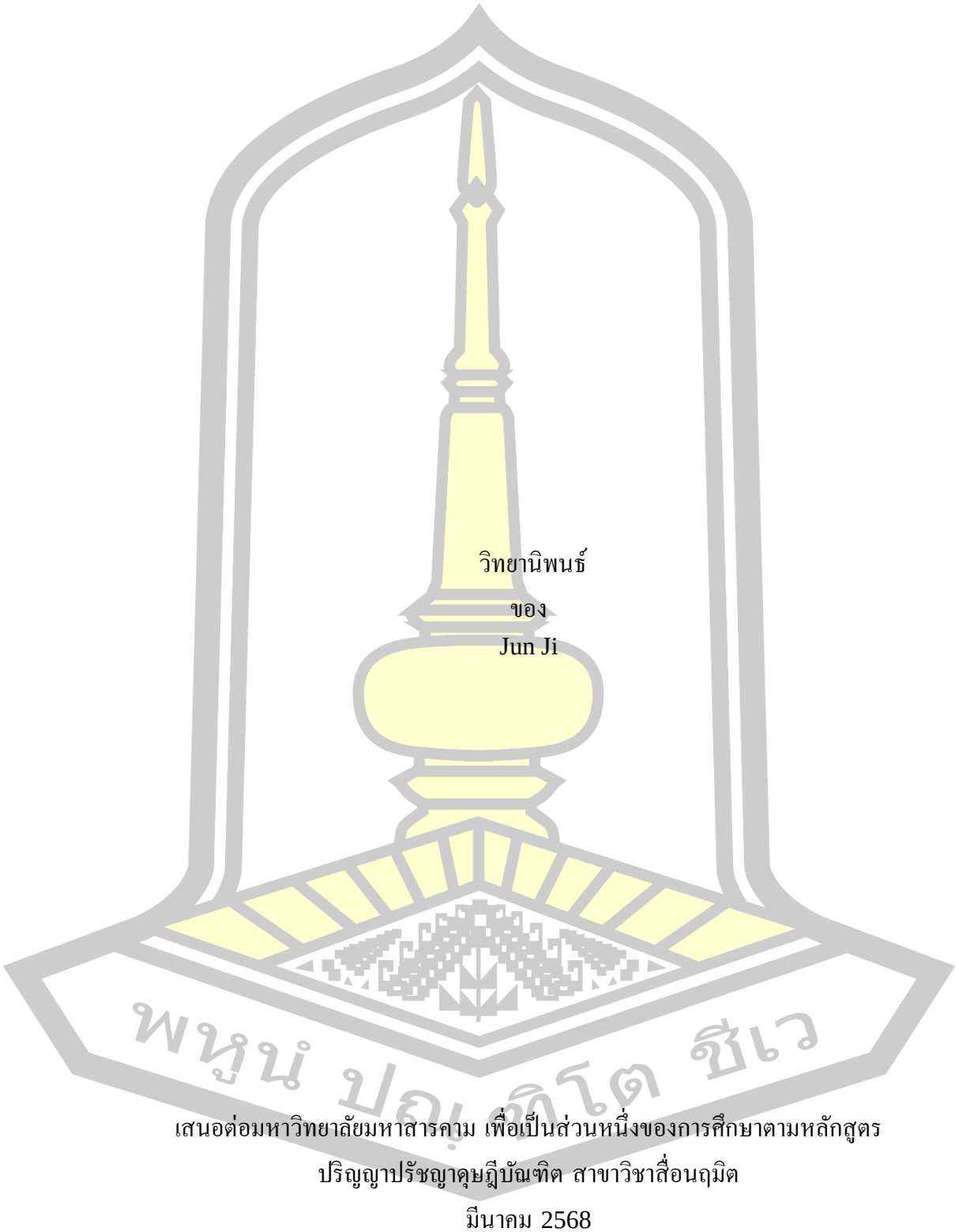
Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

Jun Ji

A Thesis Submitted in Partial Fulfillment of Requirements for
degree of Doctor of Philosophy in Creative Media
March 2025

โมเดลเสียงเพื่อสะท้อนอารมณ์เชิงบวกของผู้ชมสำหรับพอดแคสต์



วิทยานิพนธ์

ของ

Jun Ji

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาสื่อนฤมิต

มีนาคม 2568

Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

Jun Ji

A Thesis Submitted in Partial Fulfillment of Requirements

for Doctor of Philosophy (Creative Media)

March 2025

The examining committee has unanimously approved this Thesis, submitted by Mr. Jun Ji , as a partial fulfillment of the requirements for the Doctor of Philosophy Creative Media at Mahasarakham University

Examining Committee

.....Chairman
(Assoc. Prof. Nirat Soodsang , Ph.D.)

.....Advisor
(Asst. Prof. Kotchaphan Youngmee , D.F.A.)

.....Co-advisor
(Asst. Prof. Khachakrit Liamthaisong , Ph.D.)

.....Committee
(Asst. Prof. Teerayut Pengchai , Ph.D.)

.....Committee
(Assoc. Prof. Ratanachote Thienmongkol , Ph.D.)

Mahasarakham University has granted approval to accept this Thesis as a partial fulfillment of the requirements for the Doctor of Philosophy Creative Media

.....
(Assoc. Prof. Jantima Polpinij , Ph.D.)
Dean of The Faculty of Informatics

.....
(Prof. Anongrit Kangrang , Ph.D.)
Acting Dean of Graduate School

TITLE Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

AUTHOR Jun Ji

ADVISORS Assistant Professor Kotchaphan Youngmee , D.F.A.
Assistant Professor Khachakrit Liamthaisong , Ph.D.

DEGREE Doctor of Philosophy **MAJOR** Creative Media

UNIVERSITY Mahasarakham University **YEAR** 2025

ABSTRACT

This study aimed to: (1) investigate the impact of podcast hosts' voices and genders on the audience's positive emotional responses, (2) develop an audio voice model to reflect the audience's positive emotions for podcasting, (3) examine the correlation between a voice model reflecting positive emotions and the gender characteristics of podcast hosts, and (4) compare the audience's positive emotions before and after listening to a podcast. The research sample comprised 10 podcast hosts (5 male and 5 female) and 50 listeners, both recruited through voluntary sampling. The study employed a mixed-method approach, utilizing quantitative and qualitative research instruments, including: (1) Podcast Impact Evaluation Form, (2) Suitability and Acceptance Evaluation Form, (3) EEG Assessment, (4) Semi-Structured Interviews, and (5) Head-Mounted Brainwave EEG Measurement to capture real-time neural activity. Data analysis included descriptive statistics, paired t-tests, Pearson correlation, and Two-Way ANOVA.

The research findings revealed that:

1) Impact of Voices and Genders of Podcast Hosts on the Audience’s Positive Emotional Responses Analyzed how podcast hosts' vocal attributes and gender influenced the emotional responses of 50 listeners, with a majority of participants aged 18–33 years old (64%). Female hosts were preferred by 64% of listeners, whereas male hosts were selected by 26%, demonstrating a significant gender-based preference in podcast engagement. The results indicate that female hosts were more effective in evoking warmth, gratitude, and passion, while male hosts elicited stronger responses in joy, excitement, and contentment. The findings reinforce the notion that tonal and spectral variations in speech significantly influence emotional engagement, with higher pitch, rhythmic modulation, and expressive articulation favoring female hosts, whereas lower pitch, steady rhythm, and dynamic intonation benefited male hosts. EEG data further validated these responses, showing increased theta and gamma wave activity for female-hosted content, which correlated with enhanced memory encoding and emotional processing.

2) Developing a Speech Model that Reflects Audience's Positive

Emotions The Positive Emotions Speech Model (AVM) consists of six key components: (2.1) Input Audio Processing (2.2) Audio Analysis (2.3) Emotion Adaptive Voice Synthesis (2.4) Creative Voice Modulation (2.5) Output Audio Generation (2.6) Evaluation and Optimization. Refinement of the model through expert and listener input. AVM received the highest suitability ratings from five experts (mean = 4.80, standard deviation = 0.42). They demonstrated how the model effectively modulates alpha, beta, and theta band activity, aligning with EEG findings that show neural engagement in emotional processing and cognitive resonance. The results highlight the importance of real-time listener feedback integration in enhancing personalized podcast experiences and optimizing voice modulation for emotional impact.

3) The Pearson correlation analysis revealed significant relationships between podcast host gender and positive emotions. Female voices elicited stronger responses for gratitude, passion, and love, while male voices were more effective in generating joy, excitement, and contentment. The highlighted correlations between vocal traits (e.g., pitch, tone, rhythm) and EEG activity, confirming that gender-specific characteristics influenced brain regions associated with emotional processing. The data supported the gender-informed design of the (AVM), emphasizing its ability to align emotional delivery with listener preferences.

4) Comparison of Positive Emotions Before and After Listening A paired t-test compared positive emotions before and after podcast listening, revealing significant increases in joy, excitement, contentment, warmth, gratitude, and passion ($p < .05$), while hopeful remained stable. Female listeners experienced greater emotional enhancement than males, particularly in gratitude and passion. EEG data (Table 4.26, Table 4.27, Table 4.28) showed post-listening increases in theta, beta, and gamma band activities, confirming the emotional impact of the podcast on neural processing. The findings support the effectiveness of the AVM in enhancing positive emotions, demonstrating that tailored voice models can optimize emotional resonance and listener engagement in podcasting, contributing to the development of more emotionally impactful audio content.

Keyword : Audio Voice Model, Positive Emotional, Audience Perception, Podcasting, EEG Metrics

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my main advisor, Asst. Prof. Kotchaphan Youngmee, for their invaluable guidance, encouragement, and unwavering support throughout the journey of my doctoral research. Their expertise and insightful feedback have been instrumental in shaping this work. I am equally grateful to my co-advisor, Asst. Prof. Dr. Khacharit Liumthaisomg, Ph.D., whose constructive criticism and expert advice provided critical clarity and direction to my research.

I extend my sincere thanks to the examining committee members: Assoc. Prof. Dr. Nirat Soodsang, Ph.D., Asst. Prof. Dr. Teerayut Pengchai, Ph.D., and Assoc. Prof. Dr. Ratanachote Thienmongkol, Ph.D. Their thoughtful evaluations and valuable recommendations greatly enriched this dissertation and enhanced its academic rigor.

Special thanks are due to the experts and participants who contributed their time and knowledge to this study, as well as the survey respondents, whose input formed the foundation of my analysis. I would also like to express my gratitude to Mahasarakham University for its generous research funding and institutional support, which made this study possible.

I am deeply thankful to my friends, Xiaoqing Yang and Guanyun Zhao, for their encouragement, camaraderie, and unwavering belief in me. Their support during this journey has been a source of inspiration and strength.

Lastly, I wish to thank my family for their endless encouragement and understanding. Their unwavering belief in my abilities and their constant support have been my pillars of strength throughout this challenging yet fulfilling journey.

To all who contributed to the success of this research, I extend my deepest appreciation.

Jun Ji

TABLE OF CONTENTS

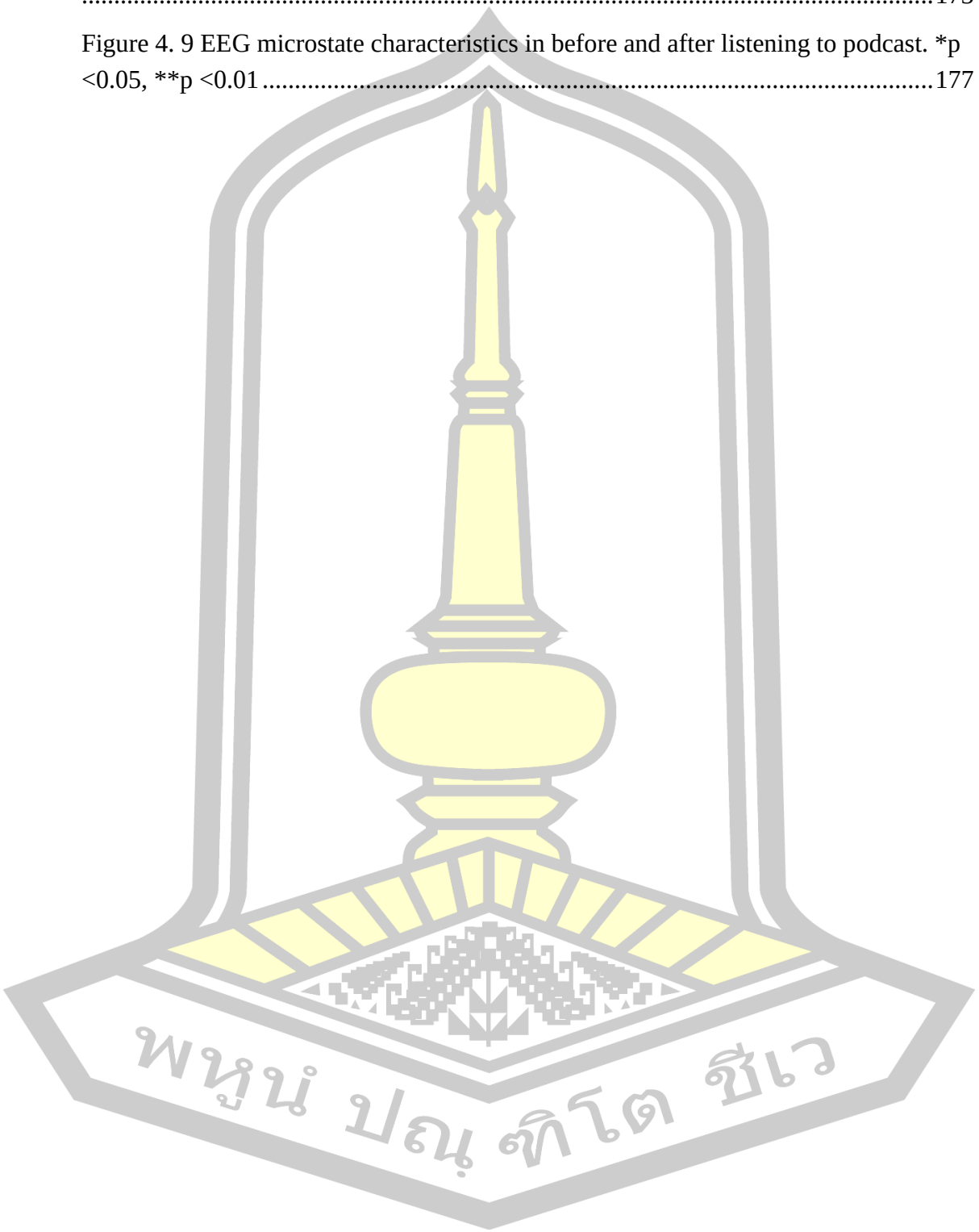
Page

LIST OF TABLES

Pages

LIST OF FIGURES

Pages



CHAPTER I

Introduction

1.1 Research Background

In today's society, economic and social competition has a significant impact on the mental health of individuals. The high level of economic and social competition is not only reflected in work and schooling, but also in social status, interpersonal relationships, and many other aspects, which bring a variety of mental health challenges to individuals. First, the high level of stress associated with economic and social competition is often one of the main triggers of anxiety and insomnia. Individuals often feel uneasy and anxious when faced with intense competition and an uncertain future, leading to decreased sleep quality. In addition, intense competition can trigger mood swings, including irritability and depression. Individuals may feel tense and uneasy from being in a constant state of competition, and even experience difficulty concentrating, which affects the efficiency of daily work and life (Duncan, Williams., 2020). In addition, individuals in competitive environments often feel lonely and isolated, lacking support and understanding, which increases the risk of mental health problems. Even worse, prolonged exposure to competitive environments may lead individuals to develop a sense of hopelessness and feel a loss of hope for the future.

The growing number of individuals facing these mental health challenges is evidenced by a wealth of research data. According to the World Health Organization (2023) statistics, as of today, more than 300 million people worldwide suffer from depression, while the number of people suffering from anxiety disorders is even higher at 260 million. Therefore, there is an urgent need for ways to alleviate mental health problems brought about by economic and social competition, which should be able to effectively help individuals cope with negative feelings such as stress and anxiety (WHO statistics, 2023).

In China, more and more people are showing interest in using podcasts to alleviate negative emotions. The White Paper on National Listening Trends released by the State Key Laboratory of Media Convergence Production Technology and

System of Xinhua News Agency (2024) shows that podcasts play an important role in enhancing people's sense of well-being. Among those who listen to audio content for stress relief and relaxation/anxiety, audio has the greatest impact on their sense of well-being, followed by those who need emotional companionship/pleasure. More than 70% of users felt that audio content improved their sense of well-being, with 45% stating that it was integrated into their lives, 29% stating that it helped them and improved their sense of well-being, 18% using it as a tool when needed and feeling rewarded, and only 8% not feeling any sense of well-being or content (Xinhua News Agency. 2024).

According to the survey, more than 60% of the audience said they have used or often use podcasts to relax, reduce stress, and enhance their emotional state. Himalaya, one of the most popular blogging platforms in China, has created more than 37,000 audiobook albums using AIGC, and the daily playback time of AIGC works has exceeded 2.5 million hours. Content) is a form of generative artificial intelligence, and according to the platform's data, the number of AIGC albums in 2023 has increased eightfold and the number of plays has increased 353% compared to 2021. These figures clearly demonstrate the prevalence of mental health issues and the potential interest in podcasting as an emotional support tool (Xinhua News Agency, 2024).

Podcasting evolves to the point where many podcast programs face several challenges in generating positive emotions through sound. While podcasts, as a form of media with emotional expression and information delivery, inspire positive emotions in listeners, the question of how to accomplish this effectively remains pressing. Current podcast content often presents problems such as poor sound quality, single program content, or a lack of emotional resonance, all of which affect listeners' emotional experiences (Jensen, P. R., 2020).

The vocal characteristics of podcast hosts play a significant role in shaping listeners' emotional engagement and overall satisfaction. Issues such as monotony in vocal delivery, lack of clarity, or an absence of emotional modulation diminish the listener's connection with the content, reducing the likelihood of sustained engagement. Research by Smith and Brown (2021) highlights that podcasts featuring hosts with dynamic and emotionally resonant voices retain listeners more effectively

and evoke stronger positive emotional responses compared to those with less engaging vocal attributes.

As the podcast market continues to grow, these challenges pose significant hurdles to attracting and retaining both creators and audiences. Addressing the limitations associated with hosts' vocal delivery, alongside other content-related issues, proves critical for fostering emotional resonance and ensuring the long-term viability of the industry. The development of an effective audio model that integrates vocal dynamics and emotional appeal emerges as a crucial solution to overcoming these barriers and enhancing the listener experience.

This research explores the human voice can be analysed and processed to better understand acoustic features such as emotional expression and intonation changes, which can more accurately reflect listeners' emotional responses and experiences. Through the study of human voice, a more intelligent and personalized audio voice model can be developed to provide a more scientific basis for the production, adjustment and optimization of podcast programs. And the proposed audio model has many potential benefits and implications. First, by addressing the current challenges faced by podcasters, the audio speech model can help improve the quality of podcast programs and enhance their appeal, thus attracting more listeners. Second, the development of audio speech modeling can provide content creators with new creative ideas and possibilities, enabling them to better convey emotions and enhance the emotional resonance of their programs. In addition, it can also help attract new podcast hosts and contribute to the diversification and prosperity of the podcast market (Jensen, P. R., 2020). Most importantly, by enhancing the overall podcast experience, the audio model is expected to provide listeners with a richer and deeper emotional experience, thereby increasing their satisfaction and loyalty.

Over the years Researchers and practitioners have explored the complex interaction between sound composition and audience emotional responses. Sound, which is an essential component of audio-visual content has a unique ability to Encourage emotions and shape the experience of the audience. Past studies have shown that manipulating sound elements, including music, ambient noise, and modulation, can have a profound influence on emotional states such as happiness, relaxation, and excitement (Juslin & Sloboda, 2010; Li et al., 2015). different It offers

valuable insights for designing audio content tailored to elicit specific emotional responses. The role of sound in visual and audio storytelling. The sound acts as an important narrator in and of itself. It enhances the emotional depth and impact of the narrative (Kassabian, 2001; Chion, 1994). Sound can strengthen character motivation. Create a realistic environment and establish an emotional connection between the audience and the content (Gorbman, 1987; Tan et al., 2020).

In recent years, there has been a growing recognition of the profound impact of sound design on audience perception and emotional engagement. Auditory elements in audiovisual content play a pivotal role in influencing how viewers understand and relate to visual narratives. Specifically, the treatment of sound elements in audio voiceover has great potential to shape viewer emotions and enhance the overall viewing experience. However, a significant gap still exists in the systematic understanding and application of sound design principles to consciously encourage positive emotional responses from viewers. To address this gap, this research endeavors to develop an innovative audio-speech model that effectively harnesses the power of sound to elicit positive emotional responses in audience perception. This model is based on a comprehensive exploration of the complex elements of sound design, including aspects of rhythm, tempo, pitch, intonation, and sound effects. By strategically integrating these elements, the model aims to create a synergistic listening experience that resonates with the audience's emotions.

The proposed study is based on psychological theories that shed light on the profound connection between sound and emotion, and scholars such as Scherer (2004) and Juslin and Västfjäll (2008) have elucidated the psychological mechanisms by which sound influences emotional states. By delving into the psychology of positive emotional responses, the present study attempts to reveal the subtle interplay between auditory stimuli and the elicitation of positive emotional states in audiences. The significance of this study lies not only in its theoretical contribution but also in its practical implications. It enables producers such as podcasters and media content creators to craft narratives that resonate deeply with their audiences and create immersive and memorable experiences.

Through an advanced audio voiceover model based on the principles of sound psychology, this study aims to tap the potential of auditory stimuli to respond to the

positive emotions of the audience. The significance of this study transcends academia and provides a practical way to enhance the emotional impact of audiovisual narratives.

1.2 Research Question

1.2.1 What is the impact of podcast hosts' voices and genders on the audience's positive emotional responses?

1.2.2 How can an audio voice model be developed to reflect the audience's positive emotions for podcasting?

1.2.3 What is the correlation between the voice model reflecting positive emotions and the gender characteristics of podcast hosts?

1.2.4 How do the audience's positive emotions differ before and after applying the voice model to podcast hosts of different genders?

1.3 Research Objectives

1.3.1 To investigate the impact of the voices and genders of podcast hosts on the audience's positive emotional responses.

1.3.2 To develop an audio voice model to reflect the audience's positive emotions for podcasting.

1.3.3 To investigate the correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts.

1.3.4 To compare of the audience’s positive emotions before and after applying the audio voice model to podcast hosts of different genders.

1.4 Research Hypothesis

1.4.1 There is a significant relationship between sound perceptual components, (pitch, volume, speech patterns, rhythm, timbre, accent, and pauses) and the audience’s positive emotions, namely Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.

1.4.2 The speaker's gender (male/female) affects the perception of positive emotions in listeners.

1.4.3 The application of the voice model (before/after) affects the perception of positive emotions in listeners.

1.4.4 There is an interaction effect between the speaker's gender and the application of the voice model on the perception of positive emotions in listeners.

1.5 Definition of Terms

1.5.1 Audio Voice Model

Audio Voice Model (AVM-PE) is a structured system designed to enhance podcast voices by reflecting positive emotions. It processes, analyzes, and adapts voice characteristics—such as pitch, rhythm, and tone—to create engaging and emotionally resonant audio. The model consists of 6 key components: input processing, emotion detection, adaptive synthesis, creative modulation, audio generation, and evaluation. Its goal is to optimize vocal delivery for better listener engagement and emotional impact.

1.5.2 Audience's Positive Emotions

The positive emotions of podcast listeners refer to the emotional responses of podcast listeners associated with 7 positive feelings: 1) Joyful, 2) Excited, 3) Contentment, 4) Loved and Warm, 5) Hopeful, 6) Grateful, and 7) Passion. These emotions are influenced by the sound characteristics of the podcast host before and after applying the sound model.

1.5.3 Podcasting

Podcasting is a digital audio broadcasting format in which spoken content is produced, disseminated, and accessed on demand. In this study, podcasting serves as a medium that uses a speech model to examine its impact on listeners' emotional responses.

1.5.4 Reflection

In this study, reflection refers to the ability of a speech model to accurately convey and enhance positive emotions in listeners through vocal modulation, which encompasses the alignment of vocal characteristics with intended emotional expressions to enhance emotional engagement.

1.5.5 Gender Characteristics of Podcasters

The gender characteristics of podcasters in this study refer to the vocal attributes and speech patterns exhibited by male and female podcast hosts selected based on their ability to convey emotions through voice modulation. Gender differences in pitch, tone, rhythm, and speech rate were analysed to determine their impact on the audience's positive emotional responses.

1.5.6 Audience

In this study, the audience refers to podcast listeners who participated in the research aged between 18–45 years. These participants actively engaged in evaluating podcast voices and provided responses regarding their emotional reactions to different podcast host vocal characteristics.

1.6 Research Scope

1.6.1 Content

The podcast content includes 7 topics that evoke positive emotions: 1) Joyful, 2) Excited, 3) Contentment, 4) Loved and Warm, 5) Hopeful, 6) Grateful, and 7) Passion. These topics were prepared for both male and female podcast hosts, with each topic being controlled to a duration of 3 minutes.

1.6.2 Positive Emotions

The study focuses on the perception of 7 types of positive emotions: 1) Joyful, 2) Excited, 3) Contentment, 4) Loved and Warm, 5) Hopeful, 6) Grateful, and 7) Passion.

1.6.3 Audio voice format

1) Format A – The 10 podcast hosts narrated the 7 positive emotion stories before applying the audio voice model.

2) Format B – The same 10 podcast hosts narrated the 7 stories using the audio voice model for voice enhancement.

1.6.4 Sound Components

This study examines 2 types of sound components:

1) Physical sound components that creatively voice modulation. These include 7 components: frequency, phase, speed, wavelength, harmonics, and sharpness.

2) Perceptual sound components that reflect and trigger positive emotions. These include 7 components: pitch, volume, speech, rhythm, timbre, accent, and pauses.

1.6.5 Population and Sample

1) Population

This study focuses on two primary groups:

1.1) Podcast Listeners – Individuals who regularly consume podcast content, representing a diverse audience in terms of age, gender, and listening preferences.

1.2) Podcast Hosts – Professional or amateur podcasters who create and present content, exhibiting a variety of speaking styles and vocal characteristics.

2) Sample

To ensure a representative sample, the study will include:

2.1) 50 podcast listeners who actively engage with podcast content at least once a week.

2.2) 10 podcast hosts (5 male and 5 female) with prior podcasting experience, ensuring a diverse range of vocal characteristics for analysis.

The sample size is determined based on previous studies in voice and emotion research, ensuring statistical reliability while maintaining feasibility in data collection and EEG analysis.

3) Sampling Method

3.1) Podcast Listeners – A stratified random sampling method will be used to balance demographic factors such as age and gender, ensuring that the sample accurately represents the general podcast audience.

3.2) Podcast Hosts – A purposive sampling technique will be employed to select podcast hosts with diverse vocal characteristics (e.g., gender, pitch, speech rate, timbre).

This sampling approach enhances the validity of the study by selecting participants relevant to the research objectives while minimizing selection bias.

1.6.6 Research Tools

1) Podcast Impact Evaluation: Evaluated the effect of the voice and gender of podcast hosts on positive emotional responses.

2) Evaluation Form: Assessed the suitability of the components of an audio voice model for reflecting audience positive emotions in podcasts.

3) EEG Assessment: Measured the audience's positive emotions before and after listening to podcasts using Electroencephalography (EEG).

4) Semi-Structured Interview: Evaluated the relationship between the audio voice model and podcast host gender on positive emotions.

1.6.7 Research Equipment

1) Voice Analysis Software: analysed the components of podcast audio.

2) The FLEX 2 Saline - 32 Channel Wireless EEG Head Cap System is a head-mounted brainwave meter designed to capture EEG signals in real time, enabling the measurement of listeners' positive emotional responses to different podcast programs. Utilizing the international 10-20 electrode placement system, this system ensures standardized and reliable brain activity recording.

Its 32-channel configuration provides comprehensive brain coverage, capturing neural activity relevant to auditory and emotional processing. By integrating EEG data with the Positive Emotions Measurement Scale, this study conducted a detailed assessment of listeners' emotional experiences, analyzing pleasure, excitement, and other affective responses to podcast content.

Figure 1. 1 FLEX 2 Saline - 32 Channel Wireless EEG Head Cap System

3) Voice Recording and Processing Software – Tools for capturing and analyzing voice data to develop the model.

1.6.8 Geographical Scope

The study focused on listeners in China who used Mandarin as their primary language.

1.6.9 Time Scope

1) The development and testing of the voice model were conducted over a one-year period.

2) Data collection, including audience feedback, was conducted over six months.

1.7 Expected Benefits

1.7.1 This research will provide an understanding of the impact of podcast hosts' voices and genders on listeners' positive emotional responses, which can be applied to designing podcast content that effectively meets audience needs.

1.7.2 It will result in a developed audio voice model capable of reflecting listeners' positive emotions for podcasting, enhancing audience engagement, and creating a more positive listening experience.

1.7.3 It will offer insights into the correlation between voice models that evoke positive emotions and the gender characteristics of podcast hosts, enabling producers to design voice styles and select hosts that align with their target audience.

1.7.4 It will provide a comparison of listeners' positive emotions before and after listening to podcasts, demonstrating the effectiveness of the developed audio voice model in increasing audience satisfaction.

CHAPTER II

Literature Review

This research on the Audio Voice Model to Reflect Audience Positive Emotion for Podcasting has reviewed the following related literature:

2.1 Podcast

2.1.1 The Meaning of Podcast

The meaning of a podcast refers to its essence or significance in terms of content, format, and purpose. It encompasses what the podcast aims to convey to its audience, the themes or topics it covers, and the overall impact it intends to have on listeners (Jensen, P. R., 2020).

2.1.2 Definition of Podcast

The term "podcast" originated as a portmanteau of Apple Computer's "iPod" and "radio". Today, it has come to refer to a type of audio program that is distributed via the Internet and can be listened to by users on various devices such as smartphones, computers, Tab-lets, etc. via streaming or downloading (Jensen, P. R., 2020).

2.1.3 Popular Podcasts in China

Popular podcasts in China cover a wide range of topics and content, and this study focuses on positive mood, upbeat types of programs. The following are some of the well-known popular Chinese podcasts including:

"Thousands of Jokes": a hilarious program featuring segments and jokes that brings joy and easy laughter to listeners. "Positive Energy Podcast": a program that delivers positive energy and motivation, sharing inspirational stories, success stories and life lessons, bringing positive energy to listeners. "Fun Voiceover": Himalaya's Fun Voiceover feature allows users to choose different segments to voiceover, whether it's classic movie lines or popular songs, allowing users to try to imitate and create, increasing the fun of learning "Here Come the Paragraphs": brings together a large number of hilarious skits, comedies and humorous stories, allowing listeners to relax in the midst of their busy lives. "Happy Hour": brings upbeat, relaxing music

and conversations, allowing listeners to unwind and enjoy the beauty of life in a pleasant atmosphere (Xinhua News Agency, 2024).

2.1.4 Popularity Podcast Analysis

The White Paper on National Listening Trends released by the State Key Laboratory of Media Convergence Production Technology and System (2024) of Xinhua News Agency shows Since 2012, audio platforms have realized a remarkable growth in the number of voices, from an initial thousand or so to more than 30,000 times more. Taking the key points in the development of the Internet as a reference, the number of voices in 2023 is 135 times higher than in 2016 and 9 times higher than in 2020. Specifically, the number of new voices increased by 893% in 2013, 245% in 2016, 160 % in 2018 and 125% in 2020 (Xinhua News Agency, 2024).

2.1.5 Types of Podcasts

Podcast content can include interviews, lectures, voiceovers, music, stories, news, and other forms of audio programs (Jensen, P. R., 2020).

Figure 2. 1 Various podcast shows (Tianhe Media, 2021)

2.1.6 Channels for Distributing and Listening to Podcasts

In China, podcast platforms are one of the most common channels for publishing and listening. These platforms include, but are not limited to, Himalaya, Dragonfly FM, Lychee FM, Apple Podcasts, Spotify, and so on.

The research in this paper focuses on Himalaya, a podcasting platform, as a major channel for podcast program distribution and listening. As one of the largest audio platforms in China, Himalaya brings together a large number of quality podcast programs covering a wide range of fields and topics. Producers can upload and publish their programs through the Himalaya platform and attract more listeners through the search and subscription features on the platform. Meanwhile, listeners can easily find and listen to the programs they are interested in on the Himalaya platform and enjoy high-quality audio content (Xinhua News Agency, 2024).

Figure 2. 2 Himalaya International Edition (Tianhe Media, 2023)

2.2 Emotions

2.2.1 The Meaning of Emotions

Emotions are the sum of an individual's subjective cognitive experiences and include a variety of general psychological and physiological states. As early as 1884, there was a definition of emotion, as well as related debates and research on emotion. However, there is still no consensus on the definition of emotion. The American psychologist, Izard (1977), started from the developmental theory and proposed that people are born with basic emotions and that complex emotions sharpness from basic emotions as the individual develops. Emotions are considered perceptible and measurable. Studied emotions by dividing complex psychological phenomena into emotions and dividing them into components, in which emotions are divided into emotional experiences, physiological arousal, and expressions (Xiting Huang, 2007). Emotions are changes that occur with different situations and needs fulfillment of

people. When people's needs in society are satisfied, positive emotions of happiness and joy are stimulated; when people's needs are suppressed and not satisfied, negative emotions such as anxiety and irritability arise.

2.2.2 The Significance of Emotions

Emotions are significant to individuals in that they provide information about their surroundings and play a key role in behavioral choices and adaptive responses. Emotions also help individuals to establish and maintain social relationships, facilitate emotional communication and the expression of feelings, and have a significant impact on an individual's mental health and well-being (Lewis, J., 2009).

2.2.3 The Nature of Emotions

Emotions are a complex psychological phenomenon involving individuals' subjective feelings and psychological responses to external stimuli. They are usually accompanied by physiological changes, such as accelerated heart rate, muscle tension, etc., and are manifested in specific emotional states, such as happiness, sadness and anger. The nature of emotions is shaped by a combination of cognitive, physiological and social factors in the individual (Ma, W., & Thompson, W. F., 2015).

2.2.4 Basic Human Emotions

Basic human emotions include joy, sadness, anger, fear, surprise and disgust. These emotions are shared by humans and play an important role in the process of biological evolution and cultural development (Lewis, J., 2009). Basic emotions have specific physiological and behavioral manifestations, and are also influenced by cultural context and individual experience.

2.2.5 Categorization of Emotions

Emotions can usually be categorized into two main categories: positive emotions and negative emotions. Positive and negative emotions form the core of the human emotional experience and have a profound impact on an individual's mental health, social interactions, and overall quality of life. Positive emotions, such as joy, excitement, contentment, and hope, are thought to be strongly linked to subjective well-being and positively contribute to immune system health and coping with life stressors, while conversely, negative emotions, including sadness, anxiety, anger, and frustration, have potentially negative effects on mental health. These emotional

responses are often accompanied by negative physiological effects and cognitive impairments, (Seligman, 2002).

Russell (1980) proposed the two-dimensional theory this theoretical model is modeled by the dimensional model of emotion is a cyclic model. Posner and Russell et al. (2005) state that the two-dimensional circular model suggests that all emotional states arise from two separate neurophysiological systems: one on potency and the other on arousal. All emotions fall in a two-dimensional space with coordinates on potency and arousal. Russell and Feldman-Barrett (1999) argued that emotions, including happiness, sadness, anger, fear, disgust, surprise, etc., can be mapped in several dimensions, such as pleasant-unpleasant, activation-inactivation. Although Russell's (1980) circular pattern theory explains the types of responses in many situations, it does not explain the overall variation in human emotional responses and it excludes the possibility of simultaneous activity of positive and negative emotions.

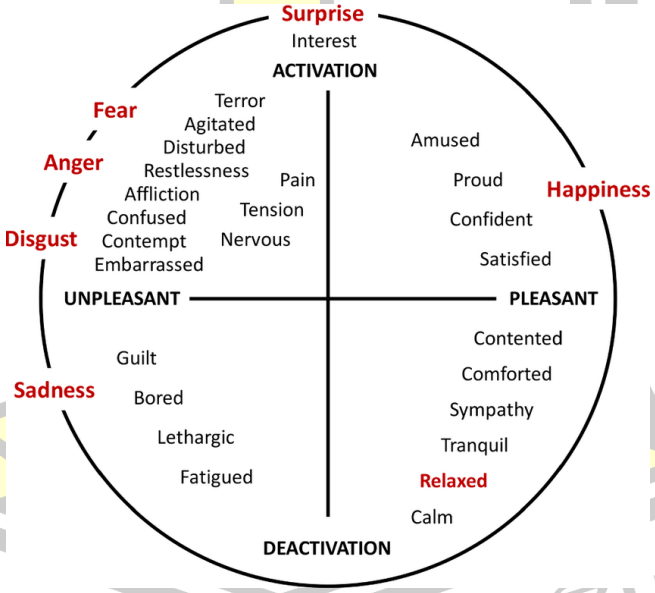


Figure 2. 3 Russel's circumplex model (Russell, 1980)

Russell argues that emotions can be categorized into two dimensions in total, which are pleasantness and intensity. Pleasantness can be categorized as pleasant and unpleasant, and intensity can be categorized as moderate and high. Pleasantness and intensity can be combined into four types: happy-high intensity is pleasure, happy-

moderate intensity is relaxation, unpleasant-moderate intensity is boredom, and unpleasant-high intensity is panic (Russell, 1980).

2.2.6 Expression of Emotions

Expression of emotions refers to the process of communicating an individual's internal emotional state through speech, facial expressions, body language, voice, and other nonverbal means. The expression of emotions can be spontaneous or a conscious attempt to communicate to others or self-expressive behavior. Such expression is usually an outward manifestation of emotional experience and can help others understand an individual's emotional state, as well as contribute to the individual's self-awareness and emotional regulation (Lewis, J., 2009). In this study, the main focus is on conveying positive emotions to the audience through words and sounds, which is closely related to the "audio voice model" in the research theme.

2.2.7 Emotional Regulation

After the 1980s, emotion regulation was separated from the field of emotion and became a relatively independent field of study. Especially in the past 20 years, emotion regulation has received active attention from physiologists and psy-chologists and has become a key area of their research.

Emotion regulation refers to the process by which individuals perceive their emotions when they occur, as well as how they are felt and expressed (Gross, 1998). Gross argues that emotion regulation can change the intensity, value, and duration of an emotion, and that emotion regulation is latent, explicit, long-term, and ever-changing, and is a dynamic process. Cole (2004) argues that emotion regulation is a correlated change that induces emotions, which can change both physiology and psychology as well as the emotions themselves. Thus, emotion regulation is the ability of individuals to perceive and control changes in their intrinsic emotions and to adapt in time to changes in their life circumstances, and is an evolving physiological and psychological phenomenon. Researchers have found that people change their emotions through subjective experiences and behaviors that reinforce positive emotions, such as happiness (Quidbach & Mikolajczak, 2010), and attenuate negative emotions, such as fear, anger, and sadness. At the same time, people lengthen or shorten the duration of their emo-tions by sharing more pleasant experiences with others, which alters the duration of emotions (Gable & Asher, 2004).

2.2.8 Emotional Development

Emotional development is a complex and dynamic process involving the gradual evolution of an individual's ability to perceive, express, and regulate emotions over the lifespan. In infancy and early childhood, individuals express emotions primarily through facial expressions, voice and body language, and gradually learn to associate these expressions with specific emotions. During childhood and adolescence, individuals' emotional cognitive abilities gradually improve, and they are able to more accurately identify and understand their own and others' emotions and learn to use appropriate language and behaviors to express them. As they enter adulthood and old age, individuals face more emotional challenges and need to cope with their emotions more effectively, and may also experience changes in areas such as physical health and social support, which may affect how they express their emotions. Thus, the development of emotions is influenced by a variety of factors, including genetic, environmental, and cultural factors, and is critical to an individual's mental health and social adjustment (Lewis, J., 2009).

2.3 Positive Emotion

2.3.1 Meaning of Positive Emotion

Positive emotions are essential components of psychological well-being, encompassing a range of uplifting feelings such as joy, excitement, contentment, love, hope, gratitude, and passion. These emotions not only enhance subjective well-being but also contribute to physical health by strengthening the immune system and increasing resilience to stress. Individuals who frequently experience positive emotions tend to cope better with challenges and exhibit higher life satisfaction (Seligman, 2002).

2.3.2 Significance of Positive Emotion

Positive emotions play a vital role in social interactions, cognitive function, and mental health. When individuals are in a positive emotional state, they are more likely to engage in meaningful social interactions, build stronger relationships, and enhance their interpersonal skills. Moreover, positive emotions improve attention, memory, and problem-solving abilities, making individuals more creative and adaptable to challenges. Thus, positive emotions not only shape an individual’s internal experience but also have a profound impact on their external relationships and cognitive performance (Seligman, 2002).

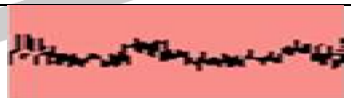
2.3.3 Types of Positive Emotion

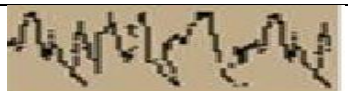
The following 7 positive emotions are widely recognized in psychological research and have been shown to influence mental and emotional well-being.

Table 2. 1 7 Positive Emotions

Positive Emotion	(Russell, 1980).	Fredrickson & Joiner (2002)	Fredrickson (2004)	Fredrickson (2013)
Joyful	✓	✓	✓	✓
Excited	✓	✓	✓	✓
Contentment	✓	✓	✓	✓
Loved and Warm	✓	✓	✓	✓
Hopeful	✓	✓	✓	✓
Gratitude	✓	✓	✓	✓
Passion	✓	✓	✓	✓

Table 2. 2 EEG of the 7 positive emotions

Emotion Category	**EEG Band**	**Brainwave State**	**Waveform Image**
Joyful	Beta (13-30 Hz	Happiness, Vitality	
Excited	High Alpha(10-13 Hz)	Excitement, Focus	
Contentment	Alpha (8-12 Hz)	Relaxation, Satisfaction	
Love	Theta (4-8 Hz)	Intimacy, Calmness	
Hopeful	Low Beta(13-20 Hz)	Positivity, Expectation	

Gratitude	Delta (1-4 Hz)	Meditation, Tranquility	
Passion	Gamma (30+ Hz)	High Excitement	

1) Joyful: Is associated with Beta wave activity (13-30 Hz), which reflects heightened mental energy, alertness, and engagement. The waveform of Beta activity is characterized by rapid oscillations with moderate amplitude, indicating an active cognitive and emotional state. This pattern suggests that joy enhances motivation and social bonding by promoting a stimulated and responsive neural state.

2) Excited: Excitement corresponds to High Alpha (10-13 Hz), signifying a balance between anticipation and engagement. The waveform shows frequent bursts of rhythmic oscillations, indicating sustained attention and heightened emotional arousal. This suggests that excitement fuels curiosity and exploratory behavior by increasing cognitive focus and emotional intensity.

3) Contentment: Contentment is linked to Alpha waves (8-12 Hz), which reflect a state of relaxation and emotional stability. The waveform of Alpha activity is smooth and consistent, with moderate amplitude and slower frequency, indicating a calm yet engaged state. This suggests that contentment fosters emotional balance and gratitude by maintaining a relaxed but conscious awareness.

4) Love: Love is primarily associated with Theta waves (4-8 Hz), which are linked to deep emotional connection and psychological comfort. The Theta waveform displays slow, synchronized oscillations with high amplitude, reflecting a meditative, immersive emotional state. This indicates that love strengthens interpersonal bonds and emotional security by facilitating relaxation and trust.

5) Hopeful: Hopeful emotions correspond to Low Beta waves (13-20 Hz), signifying a forward-thinking and optimistic mental state. The waveform shows steady, rhythmic oscillations, indicating cognitive engagement with a moderate level of excitement. This suggests that hope maintains resilience and motivation, keeping individuals focused on their goals despite challenges.

6) Gratitude: Gratitude is strongly linked to Delta waves (1-4 Hz), which promote inner peace and emotional well-being. The waveform of Alpha activity is steady and smooth, with minimal fluctuations, reflecting a tranquil and introspective state. This suggests that gratitude enhances mindfulness and social appreciation by encouraging a relaxed and positive outlook.

7) Passion: Passion is associated with Gamma waves (30+ Hz), which reflect high-intensity emotions and deep cognitive engagement. The waveform of Gamma activity is fast, dense, and highly synchronized, indicating a hyper-focused mental state. This suggests that passion drives creativity, persistence, and deep emotional involvement, inspiring individuals to pursue meaningful goals with dedication.

2.3.4 Development and Evaluation of Positive Emotion

The development of positive emotions is influenced by various factors, including early attachment experiences, emotional regulation skills, cognitive development, social interactions, self-awareness, and personality traits. Secure early relationships establish the foundation for experiencing positive emotions, while ongoing social interactions and self-reflection further shape an individual's ability to cultivate emotional well-being. Emotional regulation strategies play a crucial role in maintaining stable positive emotions, enabling individuals to cope with challenges and build resilience. Over time, people develop personalized ways of managing emotions, such as engaging in mindfulness practices, seeking social support, or participating in activities that bring fulfillment.

To measure positive emotions, researchers rely on three primary methods: self-report scales, behavioral measures, and psychophysiological assessments. Self-report scales, such as the Positive and Negative Affect Schedule (PANAS) and the Experience Sampling Method (ESM), allow individuals to evaluate their emotional experiences subjectively. Behavioral measures involve observing facial expressions, body language, and vocal tones to infer emotional states. Meanwhile, psychophysiological methods, including EEG recordings, heart rate variability, and skin conductance responses, provide objective indicators of emotional arousal and engagement (Juslin, 2008).

2.4 Sound

2.4.1 The Meaning of Sound

Sound is a mechanical wave produced by the vibration of an object that travels through air, liquids, or solids to the human auditory system and is perceived as the auditory experience of sound. In research, sound is often regarded as an important perceptual stimulus whose characteristics and meanings are important for understanding human behavior and mental processes. The meaning of sound can be direct, as in the semantic meaning in language, or indirect, as in the mood and atmosphere conveyed by background sounds in the environment (Gibbs T, 2007).

2.4.2 Perceptual Sound Components

Sound usually consists of components such as amplitude, frequency, and waveform. Amplitude determines the volume or intensity of the sound, frequency determines the tone or pitch of the sound, and waveform reflects the texture or timbre of the sound (Gibbs T, 2007). In the addition sound plays a fundamental role in human communication, perception, and emotional response. Whether in speech, music, or environmental contexts, the essential components of sound determine how it is perceived and processed. In the field of audio engineering, psychoacoustics, and communication studies, several key attributes of sound have been identified as crucial for clarity, emotional impact, and intelligibility (Howard & Angus, 2017). This literature review explores the essential components of sound, their influence on perception, and their interrelationships.

Table 2. 3 Perceptual Sound Components

No.	Perceptual Sound Components	Krumhansl (2000)	Bidelman (2013)	Tilsen & Arvaniti (2013)	Emotional Prosody (2018)	Gordon et al. (2021)	Prosody (Linguistics) (2023)	Temporal envelope and fine structure (2023)	Total Conformity.
1	Pitch	✓	✓	✓	✓	✓	✓	✓	7
2	Volume	✓	✓	✓	✓	✓	✓	✓	7
3	Speech	✓	✓	✓	✓	✓	✓	✓	7

No.	Perceptual Sound Components	Krumhansl (2000)	Bidelman (2013)	Tilsen & Arvaniti (2013)	Emotional Prosody (2018)	Gordon et al. (2021)	Prosody (Linguistics) (2023)	Temporal envelope and fine structure (2023)	Total Conformity.
4	Rhythm	✓	✓	✓	✓	✓	✓	✓	7
5	Timbre	✓	✓	✓	✓	✓	✓	✓	7
6	Accent	✓	✓	✓	✓	✓	✓	✓	7
7	Pauses	✓	✓	✓	✓	✓	✓	✓	7

The study of perceptual sound components is crucial in understanding how different auditory features influence human emotions and engagement. In the context of Audio Voice Model to Reflect Audience Positive Emotion for Podcasting, these components determine how effectively a speaker's voice can evoke and sustain emotional responses in listeners. Key Perceptual Sound Components and Their Role in Podcasting

1) Pitch

1.1 Pitch defines how high or low a voice sounds.

1.2 Higher pitches often convey excitement or urgency, while lower pitches evoke authority and calmness.

1.3 Pitch modulation enhances emotional expressiveness and listener engagement (Bidelman, 2013).

2) Volume

2.1) Volume refers to the loudness of speech, measured in decibels.

2.2) Louder voices convey confidence and passion, while softer voices create intimacy and attentiveness.

2.3) Proper volume control ensures clarity and emotional resonance (Gordon et al., 2021).

3) Speech

3.1) Speech rate determines how fast or slow a speaker talks.

3.2) Faster speech increases excitement, while a slower pace fosters reflection and comprehension.

3.3) Effective modulation of speech rate aids emotional engagement and message retention (Tilsen & Arvaniti, 2013).

4) Rhythm

4.1) Rhythm refers to the cadence and timing of speech, including pauses and stresses.

4.2) A well-structured rhythm enhances speech clarity and emotional impact.

4.3) Speech rhythm is essential for prosody, influencing perception and engagement (Prosody, 2023).

5) Timbre

5.1) Timbre is the tonal quality that makes a voice unique.

5.2) A warm and rich timbre evokes trust and emotional connection, while a harsh timbre can cause discomfort.

5.3) Timbre influences the speaker's perceived credibility and engagement (Krumhansl, 2000).

6) Accent

6.1) Accent relates to pronunciation patterns based on a speaker's regional or linguistic background.

6.2) A familiar accent can build trust and relatability, while an unfamiliar accent may affect comprehension.

6.3) Studies show that accents impact audience biases and perception of credibility (Lev-Ari & Keysar, 2010).

7) Pauses

7.1) Pauses provide structural breaks in speech, allowing listeners to process information.

7.2) Well-timed pauses enhance comprehension and create emotional impact.

7.3) Effective use of pauses increases audience engagement and retention (Beattie, 1983).

Importance in Podcasting: In podcasting, these sound components shape how listeners interpret and emotionally connect with the speaker's voice. By adjusting pitch, volume, speech rate, rhythm, timbre, accent, and pauses, podcasters can optimize their vocal delivery to evoke positive emotions like joy, excitement, contentment, gratitude, and passion.

2.4.3 Physical Sound Components

Table 2. 4 Physical Sound Components

No.	Physical Sound Components	Sahin et al. (2023)	Hu et al. (2022)	Li et al. (2021)	Fan et al. (2019)	(2019)	Zhang et al. (2023)	Kim et al. (2022)	Total Conformity.
1	Amplitude	✓	✓	✓	✓	✓	✓	✓	7
2	Frequency	✓	✓	✓	✓	✓	✓	✓	7
3	Phase	✓	✓	✓	✓	✓	✓	✓	7
4	Speed	✓	✓	✓	✓	✓	✓	✓	7
5	Wavelength	✓	✓	✓	✓	✓	✓	✓	7
6	Harmonic	✓	✓	✓	✓	✓	✓	✓	7
7	Sharpness	✓	✓	✓	✓	✓	✓	✓	7

The study of physical sound components plays a fundamental role in designing an Audio Voice Model (AVM-PE) that effectively reflects positive emotions in podcasting. These components define how sound is perceived and processed by listeners, influencing their emotional engagement and cognitive responses. The 7 key physical sound components—Amplitude, Frequency, Phase, Speed, Wavelength, Harmonic, and Sharpness—contribute significantly to the perception of vocal expression and emotion in audio content.

1) Amplitude (Loudness): Determines the intensity of sound, affecting how emotions such as excitement (higher amplitude) or calmness (lower amplitude) are conveyed.

2) Frequency (Pitch): Directly correlates with the highness or lowness of a voice, influencing how emotions like joy (higher frequency) or seriousness (lower frequency) are perceived.

3) Phase (Wave Alignment): Affects how multiple sound waves interact, impacting speech clarity and emphasis, essential for engaging podcast listeners.

4) Speed of Sound (Propagation Rate): Influences how sound waves travel through different mediums, affecting resonance and vocal clarity, crucial for immersive audio experiences.

5) Wavelength (Spatial Sound Distribution): Relates to tonal richness and the ability to create a full-bodied vocal presence, enhancing listener perception.

6) Harmonics (Tonal Complexity): Contribute to timbre, defining vocal uniqueness and emotional expressiveness in podcast speech modulation.

7) Sharpness (High-Frequency Emphasis): Refers to the prominence of high-frequency components in a sound, affecting the perceived brightness and clarity of a voice. Higher sharpness can enhance articulation and speech intelligibility but may also introduce a harsher tone if excessive, impacting listener comfort and engagement.

By incorporating these physical sound properties into AVM-PE (Audio Voice Model for Positive Emotions), the research enhances podcasting experiences by modulating vocal qualities that elicit specific emotional responses. The ability to fine-tune amplitude, pitch, timbre, and harmonics allows for a more engaging and emotionally resonant audio experience, ensuring effective communication and emotional connection with listeners.

2.4.4 Sound Production

Sound is usually produced by the vibration of an object. When an object vibrates, it produces pressure waves that travel through a medium (such as air, water, or a solid) and are interpreted as sound by the auditory system (Medvedev, Shepherd, & Hautus, 2015).

2.4.5 Characteristics of Sound

Sound has a variety of characteristics, including volume, pitch, timbre and duration. These characteristics determine the perceptual effect and expression of sound (Gibbs T, 2007).

2.4.6 Types of Sound

Sounds can be categorized into different types, including speech, music, and environmental sounds. Different types of sounds have different characteristics and meanings and have different effects on individuals (Gibbs T, 2007).

2.4.7 Sound Design

Sound design relies on sound elements, including dialog, music, and sound effects, which convey emotion, plot information, and atmosphere. Secondly, sound design takes into account how sound travels through space, including the direction, distance and echo of sound to create a realistic sound environment. Sound design also requires the use of audio equipment and technologies such as microphones, mixing consoles, sound effects libraries and stereo systems to capture, process and reproduce sound. Sound designers need to strike a balance between creativity and technicality to achieve sound design goals. Together, these elements form the core of sound design theory, adding artistry and impact to the sound of media works (Smith, John R. & Johnson, Emily A., 2023).

2.4.8 Sound Levels

Sound level is a measure of the intensity or volume of sound, usually expressed in decibels (dB). The sound level is proportional to the intensity of the sound and can be used to describe the relative loudness of the sound (Gibbs T, 2007).

2.4.9 Speed of Sound

The speed of sound is the speed at which sound waves travel through a given medium and depends on the density and elasticity coefficient of the medium. In common situations, the speed of sound in air is about 343 metres per second, but it varies in different media. The speed of sound is affected by factors such as temperature, humidity and air pressure, and increases in high temperature and high humidity environments. The speed of sound is usually higher in solids and liquids than in gases, depending on the density and elastic properties of the material (Medvedev, Shepherd, & Hautus, 2015).

2.4.10 Hearing Sound

Hearing is the ability of humans to perceive sound, receiving, processing and interpreting sound signals through the ear and auditory system. The auditory process involves the perception of external sounds, the transmission of neural signals, and the brain's interpretation and understanding of sound (Gibbs T, 2007).

2.4.11 Utilizing Sound

Sound can be used for a variety of purposes, including communication, expression of emotion, recognition of the environment, entertainment, and therapy. People use sound to communicate and express themselves through speech, music, and sound effects, as well as to improve mental health and quality of life (Medvedev, Shepherd, & Hautus, 2015).

2.4.12 Positive Emotions to Sound Components

A structured pairing of positive emotions with physical and vocal components of sound that most effectively promote them. The table also includes which elements should be increased or decreased to evoke each emotion, based on research in psychoacoustics and emotional speech processing.

Table 2. 5 Mapping Positive Emotions to Sound Components

Positive Emotion	**Increase**	**Decrease**	**Key Sound Components**
Joyful	Amplitude, Frequency, Speed, Pitch, Volume, Rhythm	Pauses, Wavelength	High amplitude, high frequency, fast rhythm, rising pitch, moderate volume (Scherer, 2003)
Excited	Amplitude, Frequency, Harmonic, Speed, Pitch, Volume, Speech	Pauses, Wavelength	Intense amplitude, high pitch, rapid speech, strong harmonics (Juslin & Laukka, 2003)
Contentment	Harmonic, Wavelength, Timbre, Pauses	Amplitude, Speed, Volume	Soft timbre, smooth wavelength, moderate harmonic presence, slow rhythm (Laukka et al., 2005)
Loved & Warm	Harmonic, Sharpness, Timbre, Speech	Amplitude, Speed	Soft harmonics, rich timbre, warm and slow speech patterns (Scherer, 2003)
Hopeful	Frequency, Wavelength,	Speed, Amplitude	Gradual rising pitch, moderate wavelength, slight emphasis in speech

	Pitch, Accent		(Juslin & Västfjäll, 2008)
Grateful	Pauses, Timbre, Sharpness, Accent	Volume, Speed	Gentle pauses, soft timbre, natural vocal flow (Scherer, 2003)
Passion	Amplitude, Frequency, Harmonic, Pitch, Volume, Speech	Pauses	Strong amplitude shifts, deep harmonics, vocal intensity, continuous speech flow (Juslin & Laukka, 2003)

This table demonstrates how the physical and vocal properties of sound influence emotional perception. By adjusting amplitude, frequency, and harmonics, along with pitch, volume, and rhythm, speakers, musicians, and content creators can enhance emotional impact in communication, storytelling, and media. Understanding these relationships allows for more effective emotional engagement in public speaking, therapy, music, and digital media.

2.5 Voice Audio

2.5.1 Meaning of Voice Audio

Voice audio refers to the recording and use of the human voice, usually in radio, television, movies, commercials, games, and other multimedia productions. This voiceover is usually done by professional voice actors who give specific sound characteristics to the recorded audio by adjusting the timbre, pitch, rhythm and range, vocal technique and other aspects of the Expression (Jeliruew, 2013).

Figure 2. 4 Voice actors and equipment (Flycase Media, 2022)

2.5.2 Components of Voice Audio

Voice audio consists of multiple components that contribute to the clarity, quality, and emotional impact of the final auditory experience. These components include human vocal recording, scripts, sound editing, mixing and post-processing, and quality control. Each element plays a distinct role in ensuring that the final audio output meets production standards and effectively conveys the intended message.

The human vocal recording process captures spoken content in a controlled environment, ensuring clarity and consistency. Scripts serve as the textual foundation for the recorded content, providing structure and direction for speech delivery. Sound editing involves modifying the recorded audio to enhance clarity, remove unwanted noise, and apply necessary adjustments to volume and tone. Once editing is complete, mixing and post-processing refine the audio by balancing multiple tracks, adjusting audio levels, and incorporating necessary effects to optimize overall quality. Finally, quality control ensures that the produced audio meets the required standards through a series of reviews, including checking the accuracy of speech articulation and overall consistency (Jeliruew, 2013).

2.5.3 Purpose of Voice Audio

The purpose of voice audio is to provide sound performance for a wide range of media productions, such as commercials, film, anime, games, education and to provide clear narration or explanation. Specifically, the purposes of audio voice include:

1) Commercials

The use of voice audio in advertising has a significant impact on guiding viewers to positive and negative emotions. Researchers Morris, Woo, Geason, and Kim have shown that emotions play a significantly larger role than cognition in predicting attitudes and behaviors of intention. This suggests that voice audio in advertisements is more likely to elicit empathy and emotional engagement from viewers through the expression of emotions, thus influencing their attitudes and purchasing behaviors (Morris et al., 2002).

Rastmi Adaval (2003) study found that consumers consider whether the emotional elements contained in advertised product information are consistent with their own emotional state when determining the weight of the product information. This implies that the emotional expression of audio voiceovers in advertisements needs to be

aligned with the emotional state of the target audience in order to better influence their decision making (Adaval, 2003).

Figure 2. 5 Voiceover personnel (Tara Strong, 2022)

2) Films

Voice audio, as a key element in the expression of emotions in movies, can lead the audience to experience positive emotions. When expressing positive emotions, the right choice of voice actors is crucial. By considering the tone, pitch, and expression of the actor, producers are able to create characterizations that reflect pleasurable, uplifting, or heartwarming emotions. (Qing Miao, 2020).

Figure 2. 6 Voice actors and voice roles (mass media, 2020)

3) Anime

In terms of animation, the application of voice audio aims to give animated characters vivid vocal expressions to enhance their personality, portray emotions and convey storylines. Through voiceover, animated characters can have unique vocal

characteristics and express rich emotions and moods, thus making it easier for viewers to empathize and emotionally connect with them. In addition, voice audio can provide clear narration and explanation for animated films, helping viewers understand the storyline and background information. The use of soundtracks and sound effects can also enhance the ambience and atmosphere of an animation, creating a more realistic and engaging viewing experience(Wujiang Zhang, 2013).

Figure 2. 7 Happy Landlord's anime voiceover (mass media, 2020)

4) Video Games

As a form of interactive entertainment, video games profoundly shape the gaming experience and guide the players' emotional experience through the sub-tle use of voice audio, thus enhancing the immersion of the game. In the game, positive emotions such as the excitement of adventure and the joy of victory, Positive emotions are further reinforced through the synthesis and processing of sound effects, and the use of audio to accentuate the climaxes of the game when the player completes quests and receives rewards (Wujiang Zhang, 2013).

Figure 2. 8 Voice audio enhances the gaming experience (Nintendo, 2022)

5) Education

Audiobook voiceover plays an important role in the field of education as a tool for vocal expression. By choosing expressive and clear voice actors, students can be presented with vivid and interesting learning content, which can improve their learning interest and comprehension. Especially in online education and distance learning, voice audio can enhance the auditory experience of learners. In addition, voice audio can be used to simulate experimental environments and improve students' experimental understanding in practical subjects. In language teaching, voice audio helps to improve students' phonological accuracy and promote language communication. Thus, voice audio brings a more vivid and diverse learning experience to education and enhances the effectiveness of learning and teaching.

Figure 2. 9 Voice Audio in Education (Tianhe Media, 2023)

2.5.4 Voice audio Expression Techniques and Characteristics

1) Pause and link

"Pausing and linking" is a technique for handling lines in the voiceover process, i.e., the judicious use of pausing and linking. Pauses give the audience time to digest the information and enhance the emotional expression, while connections maintain the coherence and rhythm of the utterance, avoiding a sense of disconnection (Wu Yu, 1999).

2) Accent of Sentence

It refers to the sounds that are emphasized in a sentence when reciting or speaking in order to express the purpose of the utterance and specific thoughts and

feelings. Stress can determine what exactly we want to tell the audience. There are many means to emphasize a word, such as adding more weight to the sound intensity and dragging the sound length(Song Zhang, 1985).

3) Tone

Tone refers to the sound form of an utterance governed by the use of specific thoughts and feelings. In voice-over creation activities, the tone of the announcer can make the audience feel whether the attitude of the script is praising or criticizing, affectionate or sorrowful, serious or solemn(Song Zhang, 1989).

4) Rhythm

Rhythm is a form of sound processing that refers to the movement of articulated speech that goes back and forth. In addition to the tone of voice, the emotional color of a text also needs to be matched by the rhythm. Determine the pace of speech when interpreting a voiceover text and determine whether the style of the content is light and happy or low and sad (Baihua Zong, 1999).

5) Timbre and Range

Timbre is the texture or character of a sound and is one of the main features that distinguishes different sounds. It depends on the harmonic structure in the sound, giving each sound its unique texture. Timbre can be described as soft, clear, rough, etc. and is one of the individual characteristics of sound. Timbre is the frequency range of a sound, usually ranging from bass to treble. Different vocals, instruments, or sound effects may have different ranges of range (Shuguang Peng, 2014).

6) Scenario Reproduction

It means that the voiceover artist, based on the materials provided by the voiceover script, makes the people, events, and situations in the script continuously emerge in his/her mind, forming a continuous moving picture, and continuously triggering pertinent emotions and stimulating a strong desire for voiceover (Minggang Chen, 2014).

2.5.5 Analyzing the Audience and Opportunities

First of all, audience analysis is a meticulous study of the target audience, including age, gender, cultural background, interests and hobbies. This kind of analysis can help voiceover teams better understand the preferences and needs of the audience, so as to develop more accurate voiceover strategies. Secondly, with the continuous

development of digital media and Internet technology, people's demand for audio content is increasing, which brings a broad market space for the voiceover industry (Xingchen Dong, 2018).

2.5.6 Preparing the Content

Preparing content is a comprehensive process involving goal setting, information gathering, structural organisation, creation, editing, design, feedback and distribution. First, the purpose and target audience of the content is clarified in order to develop the appropriate style and depth. Next, conduct in-depth research and collect relevant data and resources to ensure the accuracy of the information. Then, organise the information in a logical sequence to create a clear outline. Write or produce content based on this, and edit and proofread meticulously to ensure it is error-free. For visually presented content, focus on design and layout to enhance messaging. Get feedback and make necessary revisions before publishing. Finally, choose the right platform to publish the content to ensure that it is effectively communicated to the target audience (Xingchen Dong, 2018).

2.5.7 Steps in Voice Audio

Voice audio production is a structured process involving defining project requirements, script adaptation, voice casting, recording, editing, and quality control. Murf Studio (2023) and Tong (2023) emphasize the importance of identifying project-specific requirements, such as tone, style, and technical specifications, to maintain consistency throughout production. Sun (2020) and Li (2021) highlight script adaptation as a crucial step, particularly for multilingual content, ensuring emotional and informational accuracy across different audiences. Ji (2022) and Blanc (2022) argue that selecting appropriate voice actors enhances clarity, emotional depth, and audience engagement, as a well-matched voice strengthens listener connection. During recording, Li (2022) underscores the significance of voice modulation and pitch control in shaping the listening experience. Murf Studio (2023) and Ji (2022) further emphasize the necessity of multiple takes to achieve optimal articulation while minimizing background noise interference. Editing and mixing, as noted by Tong (2023), involve balancing volume, refining speech clarity, and eliminating distortions to ensure a polished final product.Quality control is essential before release, with Sun (2020) stressing the need for verifying pronunciation, emotional delivery, and sound

consistency. Li (2022) suggests that strategically planning the release date optimizes audience engagement, ensuring that content reaches listeners at peak times.

In summary, existing literature outlines a systematic approach to voice audio production, emphasizing both technical precision and audience alignment. However, current research lacks a thorough examination of how podcast hosts' tone, pitch, and delivery style influence listeners' emotional responses. This study aims to bridge that gap by integrating EEG-based emotional response analysis with voice design principles, contributing to the development of an audio voice model that enhances positive emotional engagement in podcasting.

2.5.8 Factors Contributing to Successful Voice Audio

Successful voice audio depends on a combination of several key factors. First, the voiceover artist must keep the language natural and loose and avoid over-exaggerating it in the studio. Secondly, the voiceover artist needs to have excellent voice control and be able to adjust pitch, volume and tempo according to the character and situation to ensure expressive and emotional communication. In addition, professional recording equipment and environment are also key factors. Finally, sound processing, mixing and editing in the post-production stage are crucial to improve the overall quality and emotional depth of the voiceover (Yuanyuan Feng, 2013).

2.5.9 Relationship Between Voice Audio and Listening

The relationship between voice audio and listening can be explored on two theoretical levels:

1) Listening Style

Behne (1997) found that there are differences in the way individuals listen to voice audio. Some listeners pay attention to the melodic lines of voice audio, others Reflect nostalgic memories of past events when stimulated by voice audio, and others pay attention to the emotional content of a voice audio piece and the ideas it expresses.

2) Symbolic Communication and Meaning Construction

Under the perspective of semiotics, sound and language in voiceover are symbols that convey meaning. Through the symbol system of sound, the voiceover actor transforms abstract textual information into concrete sound symbols, which are decoded in the listener's mind, thus realizing the transmission of information and the construction of meaning (Papotti, D., 2013).

2.5.10 Evaluating Voice Audio

The assessment of voice audio aims to ensure that the dubbed content is of the best quality and reproduction level. In voiceover practice, accurately assessing and determining pitch and emotion ensures that the voiceover is timed and paced in the same frequency, avoiding unnecessarily rushed delivery. At the same time, a robust review process and quality assurance checks are essential to identify and correct errors before final release. In addition, for legal and ethical reasons, the voice-over production process needs to comply with the regulatory standards of the particular region and be copyright compliant. (Audio Engineering Society, or AES, 2020)

Figure 2. 10 AES NY 2023 Celebrates 75th Anniversary (Kelly Reynolds, 2020)

2.6 Brain Waves

2.6.1 Meaning of Brain Waves

A certain sum of post-synaptic voltages generated by vertebrate cells in the cortex when information is exchanged between nerve cells in the brain, providing information about the brain's bioelectrical activity (Hossan, A., & Chowdhury, A. M., 2016).

2.6.2 Benefits of Brainwave Entrainment

Brainwave entrainment is a process in which the brain is guided to produce brainwaves of a specific frequency by means of external stimuli or specific tech-niques. Help to regulate the frequency of electrical activity in the brain, particu-larly by

inducing the production of alpha waves (8-14 Hz), and also improve an individual's positive mood. In an alpha wave state, brain activity slows down, the body relaxes, and the heart rate and breathing calm down, people are more likely to feel a sense of inner peace and contentment, which promotes the creation and enhancement of positive emotions (Mikko, Viinikainen et al., 2022).

2.6.3 Types and Characteristics of Brain Waves

Brain waves are typically classified into alpha (α), beta (β), theta (θ), delta (δ), and gamma (γ) waves, each characterized by distinct frequency, amplitude, and waveform morphology. Frequency, measured in hertz (Hz), represents the number of oscillations per second, while amplitude, measured in microvolts (μV), reflects the intensity and synchronization of neuronal activity. The morphology of brain waves varies depending on neural states and cognitive functions. As shown in Figure 2.11 Morphology of various EEG waveforms (Tianhe Media, 2020), different brain waves exhibit unique waveform characteristics that correspond to various cognitive and emotional states (Sascha et al., 2020).

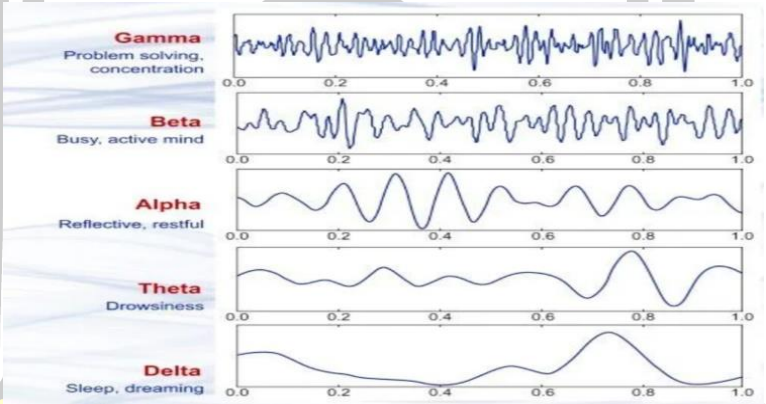


Figure 2. 11 Morphology of various EEG waveforms (Tianhe Media, 2020)

Alpha waves (8–12 Hz) are moderate-amplitude rhythms associated with relaxed, wakeful states and are prominent during moments of calm focus. Beta waves (12–30 Hz) have lower amplitudes and are linked to active thinking, problem-solving, and cognitive engagement. Theta waves (4–8 Hz), with mid-range amplitude, appear during deep relaxation and meditative states, often associated with memory processing. Delta waves (0.5–4 Hz) are high-amplitude, slow-frequency waves observed in deep sleep, critical for physiological restoration. Gamma waves (30–45

Hz), with the highest frequency and lowest amplitude, are associated with high-level cognitive functions such as perception, problem-solving, and consciousness.

2.6.4 Brainwave States and Benefits

The state of brain waves reflects the pattern of electrical activity of the brain in different cognitive and emotional states. In the alert state, the brain shows a mixture of high-frequency β and α waves, manifesting wakefulness and concentration; in the relaxed state, the α waves are more prominent, reflecting physical and mental relaxation and a calm state of mind; the sleeping state is ac-companied by the appearance of θ and δ waves, representing the entry into the sleep process (Mikko, Viinikainen et al., 2022).

Benefits of each type of brainwave: α waves occur to help relax, reduce stress, and improve concentration, as well as to stimulate creative thinking. β waves increase alertness and attention, and help to enhance active thinking and cognitive functioning. θ waves are associated with deep relaxation and meditative states, and help to bring about inner calmness and enhance creativity. δ waves are usually found in deep sleep and help promote body and brain repair and enhance memory and learning processes. γ waves are associated with perception, attention and cognitive flexibility and can enhance perceptual abilities and the ability to solve complex problems(Sascha, Frühholz et al., 2020).

2.6.5 Brain Waves and Emotions

There is a close relationship between brain waves and emotions, and different emotional states cause the brain to produce different types and frequencies of brain waves. Emotional fluctuations can directly affect the pattern of electrical activity in the brain, and the emotional state of an individual can be indirectly inferred by monitoring the pattern of brain wave activity. Emotional states can influence the generation of brain waves, while patterns of brain waves can also influence an individual's emotional experience. This bidirectional effect implies a complex interplay between emotion and brain activity (Mikko, Viinikainen et al., 2022).

2.6.6 Using Sound to Modulate Brain Waves for Emotional Regulation

The use of sound to modulate brain waves for mood regulation is a method of achieving mood regulation through sound stimulation at specific frequencies and rhythms in order to influence the brain's pattern of electrical activity. This method

involves the use of sound to produce sound waves of a specific frequency, which in turn stimulates the brain to produce corresponding brain waves. By regulating the activity of the brain waves, mood regulation effects such as relaxation, anxiety reduction, and increased concentration can be promoted (Hossan, A., & Chowdhury, A. M., 2016).

2.6.7 Brainwave Measurement Tool: Electroencephalogram (EEG)

This is a tool used to measure the electrical activity of the brain, known as an electroencephalogram (EEG). EEG records the electrical activity of neurons in the cerebral cortex by placing electrodes on the scalp. This tool is used to capture and record various electrical signals in the brain in order to further analyze and understand the functional state of the brain (Hossan, A., & Chowdhury, A. M., 2016).

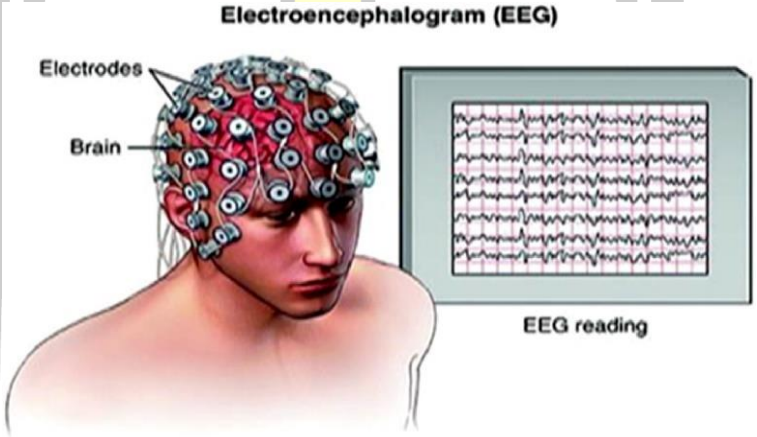


Figure 2. 12 Brain (EEG) image presentation (Tianhe Media, 2020)

2.7 Perception

2.7.1 Meaning of Perception

Perception is the process of perceiving and cognizing the external environment through the sensory organs and involves the ability to receive, process and understand external stimuli and to organize and interpret information (Milliman, R. E., 1986).

2.7.2 Importance of Perception

The importance of perception lies in the fact that it is the key ability of individuals to acquire and understand information about the external world. Through perception, individuals are able to realize the changes and characteristics of their surroundings and thus make adaptive responses and decisions. The different levels

and modalities of perception directly affect an individual's behavior, emotions, and socialize (Milliman, R. E., 1986).

2.7.3 Sensory Input in Perception

Sensory input refers to information received through the senses, in-cluding visual, auditory, tactile, olfactory, and gustatory stimuli. Visual input includes light and color, which are processed by the visual system; auditory in-put includes sound waves, which are processed by the auditory system; tactile input includes touch and pressure, which are processed by the somatosensory system; olfactory input includes odors, which are processed by the olfactory sys-tem; and gustatory input includes taste, which is processed by the gustatory sys-tem (Pannese, A., 2012).

2.7.4 Processes of Perception

The perceptual process is the process of receiving, processing and interpreting sensory input from the environment. This involves the reception of sensory input, its transmission to the brain, its processing and interpretation, and ultimately leads to the perception and understanding of the external environment. The perceptual process involves a wider range of cognitive activities, including the processing and interpretation of sensory input (Milliman, R. E., 1986).

2.7.5 Stages of the Perception Process

The perception process involves multiple stages, beginning with sensory input, where stimuli such as sound, light, and touch are detected by sensory organs and transmitted to the brain via neural pathways. The brain then analyzes and interprets this information by integrating past experiences and cognitive patterns to form a structured understanding of the environment. This interpretation leads to decision-making and behavioral responses, enabling individuals to react appropriately to sensory stimuli (Milliman, R. E., 1986).

2.7.6 Mechanism of Perception

Perceptual mechanisms refer to the process of how humans or other organisms receive, process, and understand sensory information from the external environment. It involves the reception of sensory input, its transmission to the brain, its processing and interpretation in the brain, and ultimately the formation of perception and understanding of the external environment (Pannese, A., 2012).

2.7.7 Components of Perception

Components of Perception refers to the various essential elements or components of the perceptual process that form the basis of the perceptual process. These components include sensory organs, the nervous system, perceptual selection mechanisms, perceptual processing and encoding, perceptual integration and understanding, and perceptual feedback (Milliman, R. E., 1986).

2.7.8 Factors Influencing Perception

Factors affecting perception include physiology, environment, attention and awareness. Physiological factors cover the individual's physiological state, the health of the sensory organs, and age and gender. Environmental factors include the external environment such as light, sound, odor, and temperature. Attention and consciousness determine what information is attended to and processed. These factors interact to shape the way an individual perceives and understands the external environment (Pannese, A., 2012).

2.7.9 Assessment of Perception

Perceptual assessment is a systematic process designed to assess an individual's ability to perceive and understand the external environment. This assessment process involves collecting, analyzing, and interpreting information about sensory input, processing, perceptual outcomes, and related behavioral responses (Pannese, A., 2012).

2.7.10 The audience's positive emotions are reflected by EEG

The study of brain function in emotional processing has identified key regions responsible for cognitive and affective responses. The frontal lobes (F7-F8, FC5-FC6) are central to decision-making, motivation, and emotional regulation, while the temporal and parietal lobes (T7-T8, P7-P8) process auditory input and integrate sensory experiences with emotional memory. Additionally, the occipital lobes (O1-O2), primarily associated with visual processing, contribute to multimodal emotional perception. Understanding the role of these brain regions in emotional engagement with auditory content, such as podcast listening, provides valuable insights into how vocal attributes influence listener experience. EEG has emerged as an effective tool for studying neural activity related to emotional states. For instance, Baldassano et al. (2017) found that narrative-driven content enhances theta and alpha power, particularly in temporal and parietal regions, suggesting relaxed attentiveness.

Similarly, Sachs et al. (2020) highlighted that musical elements in audio content stimulate the mesolimbic pathway, a key component of the brain's reward system, enhancing engagement and emotional arousal. These findings underscore the importance of sound design and voice modulation in shaping listeners' emotional responses. Figure 2.13 illustrates the anatomical structure of these regions and their functions in emotional and cognitive processing.

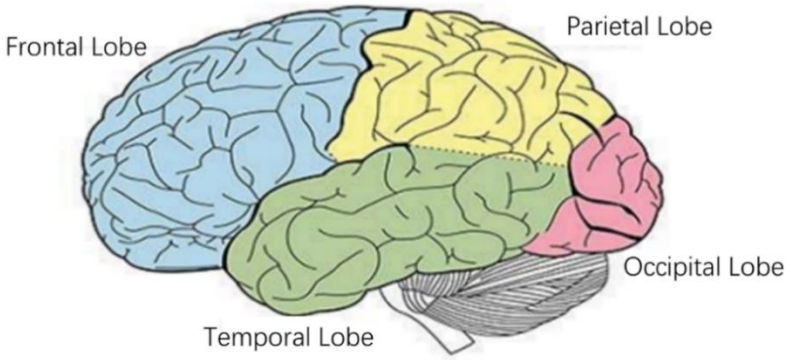


Figure 2. 13 Structure of human brain and its lobes

To further analyze emotional responses to different podcast features, EEG studies have examined relative power spectra across key frequency bands, including delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–45 Hz). As shown in Figure 2.14, EEG signals recorded from the frontal, temporal, parietal, and occipital regions reveal distinct neural patterns associated with joy, excitement, hope, gratitude, and passion. Davidson and Fox (1982) reported that increased theta and delta power in the frontal regions (F7-F8, FC5-FC6) suggests stronger emotional engagement and cognitive integration, particularly in response to joy and gratitude. Similarly, Klimesch (1999) found that higher theta power in the temporal and parietal regions (T7-T8, P7-P8) is linked to excitement and passion, indicating heightened sensory and emotional processing during podcast listening. Meanwhile, Laufs et al. (2003) observed that the occipital regions (O1-O2), typically associated with visual perception, exhibit variations in alpha and beta power, suggesting that mental imagery may play a role in emotional engagement, particularly for hope and excitement. The plots in Figure 2.14 illustrate these frequency

distributions, with different colored overlays representing emotional conditions, while adjacent bar graphs quantify mean power and standard deviations across frequency bands.

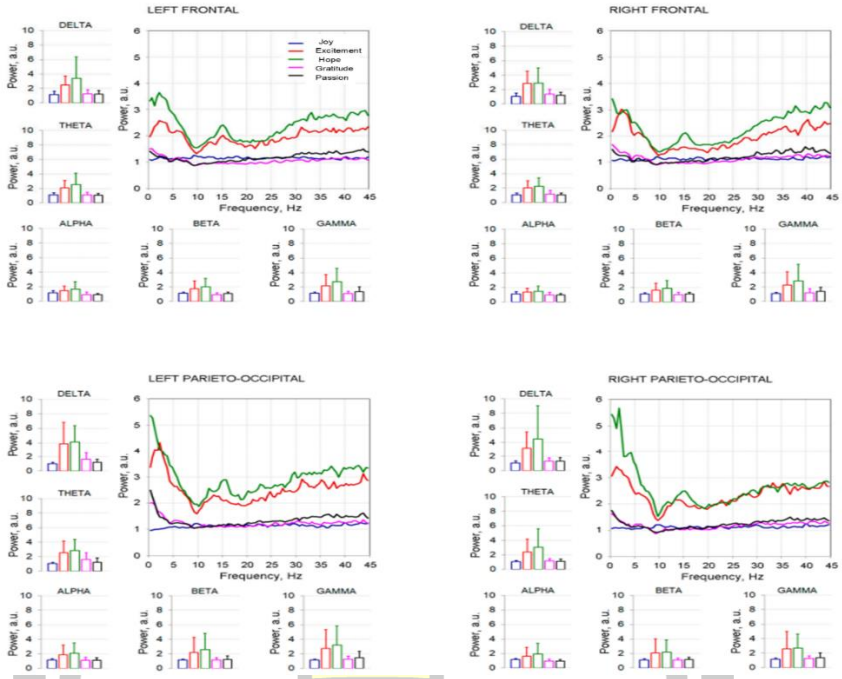


Figure 2. 14 Plots of relative power spectrums obtained

In summary, EEG-based research provides critical evidence on how vocal attributes in podcasts influence neural activity and emotional engagement. Increased theta and delta power in the frontal and temporal regions reflects stronger listener resonance with emotions such as joy, excitement, and gratitude. These findings underscore the role of voice modulation, pitch, and delivery style in shaping auditory experiences, highlighting the need for strategic vocal design in podcasting. While previous studies have examined music and narrative elements in enhancing engagement, there remains a research gap in understanding how podcast hosts' vocal attributes and gender impact EEG-based emotional responses. This study aims to address this gap by analyzing the neural mechanisms underlying vocal-driven emotional engagement, contributing to the development of an optimized audio voice model for podcasting applications.

2.8 Communication

2.8.1 Meaning of Communication

Communication refers to the process of exchanging and transmitting information, thoughts, feelings or ideas through speech, writing and body language. In communication, the sender of a message conveys the message to the receiver through various means, and the receiver receives, understands and interprets the message and responds accordingly to the sender through feedback or response (Shuhua Zhang, 2002).

2.8.2 Objectives of Communication

Communication purposes are the specific goals and intentions pursued by individuals or groups when exchanging information. These purposes usually involve the transmission of information, the expression of emotion, influence and persuasion, relationship building and maintenance, conflict resolution, action coordination, self-expression and fulfilment, social identity, and education and learning. The purposes of communication may vary in different situations, cultural contexts and individual needs, but in general they aim to promote understanding, achieve consensus, fulfil individual or collective goals, and enhance the quality of social connections and interactions (Maria, T., & Russo, T., 2010).

2.8.3 Benefits of Communication

The benefits of communication refer to the positive outcomes obtained through effective communication that promote accurate understanding of information, enhance interpersonal relationships, improve the quality of decision-making, facilitate problem solving, enhance teamwork, improve individual and organisational effectiveness and support social cohesion. The benefits of communication are seen not only in immediate interactions but also in long-term social and organisational development. Effective communication reduces misunderstandings and conflicts, increases productivity, enhances innovation and promotes inclusiveness and understanding in diverse environments (Maria, T., & Russo, T., 2010).

2.8.4 Types of Communication

The types of communication can be divided into four categories:

1) Internal communication

This refers to communication within an individual, i.e., the process by which a person engages in dialogue and communication with his or her inner self. This type of communication may involve internal activities such as thinking, feeling, decision making etc. of an individual (Shuhua Zhang, 2002).

2) Interpersonal communication

This refers to the direct communication and interaction that takes place between two or more individuals, such as face-to-face dialogues, telephone exchanges, e-mails, and so on. This form of communication is commonly found in everyday life and work situations (Shuhua Zhang, 2002).

3) Group communication

This refers to communication between individuals and groups, such as the interaction between teachers and students in the classroom, leaders' speeches in assemblies, and so on. In this type of communication, information is transferred, understood and shared between individuals and collectives to achieve a specific goal or task (Shuhua Zhang, 2002).

4) Mass communication

This is a form of information dissemination through mass media such as radio and television. Mass communication includes media in the form of magazines, pamphlets, leaflets, posters, etc., which are capable of delivering the same message to a wide range of audiences at the same time or in parallel with a wide range of reach and impact (Shuhua Zhang, 2002).

2.8.5 Principles of Communication

The principles of communication are a set of basic rules designed to enhance the efficiency and effectiveness of communication. These principles include clarity, which requires messages to be clear and free of ambiguity; conciseness, which emphasis the directness and conciseness of messages; completeness, which ensures that comprehensive information is conveyed; timeliness, which highlights the currency of the message; consistency, which emphasis the coherence of messages and action ; appropriateness, which focuses on matching the mode of communication with the environment; feedback, which advocates the establishment of an effective mechanism for the confirmation of messages; respect, which advocates respect for all parties in communication; integrity, which emphasis the truthfulness and ethicality of messages;

and purpose. integrity, which emphasis the truthfulness and ethical nature of information; and purpose, which ensures that communication activities are clearly goal-oriented (Maria, T., & Russo, T., 2010).

2.8.6 Forms of Communication

Oral communication relies on the direct expression of words, including face-to-face dialogues and long-distance calls, and it emphasizes immediacy and interactivity. Written communication, on the other hand, records information through words, such as letters, reports and e-mails, and it allows for the persistence and wide distribution of information. Non-verbal communication conveys messages through non-written means such as body language and facial expressions, and it plays a key role in enhancing the expressive power of verbal messages. Visual communication utilis visual elements such as graphics and images, which provide intuition and intensity to the message. Group communication emphasis collaboration and discussion within a team, while cross-cultural communication requires a deep understanding of communication styles in different cultural contexts (Maria, T., & Russo, T., 2010).

2.8.7 Communication Process

The communication process is a continuous interaction in which information is transmitted, understood and fed back between at least two participants. The process begins at the source of the message, which translates the intention into a transferable message. The message is then encoded and transmitted through selected channels, and the receiver decodes the message as it is received, transforming it into personal understanding. To ensure that the message is correctly understood, the receiver provides feedback, and this feedback loop is critical to the success of the communication. Throughout the process, noise can interfere with the clear delivery and reception of the message, so effective communication strategies need to consider how to minimise the impact of noise. Ultimately, the aim of communication is to achieve the desired effect, such as influencing the behaviour or attitude of the receiver (Maria, T., & Russo, T., 2010).

2.8.8 Principles of Receiving Information

The principles of message reception emphasis the importance of ensuring that messages are correctly decoded and understood during the communication process. These principles include keeping an open mind to accommodate multiple viewpoints,

focusing attention to capture details of the message, actively participating to facilitate in-depth discussion of the message, applying critical thinking to assess the quality and credibility of the message, maintaining emotional neutrality to ensure objective understanding, demonstrating cultural sensitivity to adapt to the delivery of the message in a different cultural context, providing effective feedback to facilitate two-way communication, and integrating old and new knowledge to build a comprehensive cognitive system(Watzlawick, P. et al., 1975).

2.8.9 Communication Models

The communication model refers to a systematic framework that describes the transfer of information between a sender and a receiver, and it covers the key aspects of encoding, transmission, decoding, and feedback of information. This model emphasises the bidirectional and interactive nature of communication, in which the source of information transforms intentions into transmittable messages that are passed on to the receiver through specific channels. The receiver then decodes this information to understand its meaning and may provide feedback, a process that can be affected by noise. Communication patterns involve not only the direct transfer of information, but also environmental factors such as culture, social structure and situational context, which together shape the complexity and diversity of communication. Effective communication patterns need to take these elements into account to ensure that messages are accurately delivered and received in order to fulfil the purpose of communication (Watzlawick, P. et al., 1975).

2.8.10 Efficiency of Communication

Communication efficiency refers to the ability of a message to be quickly and accurately delivered, understood and implemented with minimal investment of time, effort and resources in a given communication process. Achieving communication efficiency requires the sender of the message to clarify the communication objectives, select appropriate communication channels, code the message effectively, and establish a positive feedback mechanism. At the same time, cultural and situational factors need to be taken into account in the communication process to reduce distractions, improve the communication skills of participants, and continuously optimise communication strategies (Maria, T., & Russo, T., 2010).

2.9 Research Related

2.9.1 The Psychology of Sound and Emotional Resonance

The psychology of sound examines how auditory stimuli, such as voice, music, and sound effects, influence human emotions and cognitive states. Siu-Lan Tan et al. (2020) explored various aspects of sound perception, including acoustics, auditory processing, and the neurological responses elicited by auditory input. Their research highlights that sound perception is a multidimensional process involving both physiological and psychological factors. Penny Bergman et al. (2021) expanded on this idea, demonstrating that the categorization of sound perception is influenced by both perceptual and emotional dimensions. Their psychoacoustic experiments revealed that emotional responses to sound are shaped by a combination of learned associations and innate auditory processing mechanisms.

Further investigations by Erkin Asutay et al. (2019) and Anton Killin et al. (2020) emphasize the role of auditory attention in emotional processing. Asutay et al. (2019) used behavioral and neuroimaging studies to show that emotional significance modulates auditory perception, meaning that individuals focus more on sounds that hold affective relevance. Killin et al. (2020) introduced the concept of polysemic analysis, arguing that sound is interpreted differently depending on contextual and cultural factors. Additionally, Asutay et al. (2022) examined how the brain processes emotional rhythms, finding that lateral temporal lobes, particularly in the right hemisphere, are crucial for recognizing emotional intonations in speech.

In summary, the psychology of sound is a multidisciplinary field that integrates neuroscience, musicology, and cognitive psychology to understand how auditory elements evoke and regulate emotions. While existing studies provide valuable insights into sound perception, there is still a gap in understanding how voice-specific attributes, such as tone and rhythm, shape positive emotional engagement in digital media like podcasts.

2.9.2 The Effect of Audio Voice on Emotional Responses

Audio voice has been shown to influence emotional experiences across different contexts, including entertainment, mental health, and cognitive performance. Jung-Ki Nam et al. (2023) found that listening to audio content while performing simple repetitive tasks enhances task performance, reduces fatigue, and increases satisfaction,

enjoyment, and immersion. Similarly, Sifan Wu et al. (2023) used deep learning techniques to classify human speech emotions, demonstrating that auditory components, including pitch and tone, significantly impact affective states. Their findings suggest that speech duration and modulation are crucial in determining how emotions are conveyed through voice.

The influence of audio voice extends beyond speech. Sogand Mohammadi et al. (2023) examined emotional responses to audio components in movies, using neural networks to assess the interaction between sound, speech, and background music. Their results highlight that sound plays a dominant role in predicting valence, or the pleasantness of an emotional experience. Additionally, Gonzalo Iturregui-Gallardo et al. (2019) investigated how voiceovers and audio subtitles impact blind and partially sighted individuals, revealing that the mode of audio presentation affects emotional engagement. Duncan Williams (2020) studied video game soundtracks, showing that timbral characteristics of voice and sound effects shape listeners' emotional responses in interactive environments.

In summary, these studies confirm that voice characteristics, including pitch, rhythm, and tone, are critical in shaping emotional experiences. However, most research focuses on individual elements rather than a holistic voice model. This study seeks to bridge this gap by examining how podcast hosts’ voices influence positive emotional responses and how an optimized voice model can enhance listener engagement.

2.9.3 Sound Design and the Psychology of Positive Emotions

Sound design plays a fundamental role in eliciting emotions and creating immersive experiences. Cuadrado et al. (2020) explored how sound spatialization and audio mixing techniques impact children’s emotional engagement with audiobooks. Their research found that 3D audio significantly enhances positive emotions and mental imagery, suggesting that spatial audio can be an effective tool for emotional storytelling. Similarly, Kathryn Sharpe Wessling (2018) applied positive psychology principles to sound design, proposing that certain sound characteristics are contagious and can influence emotional well-being.

Niklas Rönnberg et al. (2016) emphasized the importance of sound in visualization design, arguing that emotional qualities in sound contribute to stronger

cognitive and affective responses. Their findings align with MA Chundong (2022), who examined the role of cognitive psychology in sound design and suggested that listeners' emotional needs must be considered in audio product development. Additionally, Yi-Yoen Kim et al. (2022) investigated interactive audio design and found that sound-based emotional stimuli can guide user interactions in digital media.

In summary, research on sound design and positive emotions underscores the powerful influence of auditory cues on emotional engagement. While prior studies explore the impact of sound on specific user groups, there is still a need for research focusing on how different vocal attributes in podcasting affect sustained emotional responses.

2.9.4 Head-Mounted Brainwave Devices for Mood Monitoring

Advancements in neuroscience have led to the development of head-mounted EEG devices that provide real-time monitoring of brain activity. Cai Zuobin (2019) introduced a wearable EEG instrument capable of detecting neural signals with high accuracy while maintaining user comfort. Similarly, Gary E. Strangman et al. (2020) developed multimodal wearable brain imaging systems that integrate physiological monitoring to assess cognitive and emotional states.

Further innovations by Cui Xingran et al. (2021) include EEG headbands designed for real-time depressive emotion monitoring. These devices use advanced signal processing techniques to detect neural patterns associated with mood states. Michael N. Sawka et al. (2022) emphasized the potential of wearable physiological sensors in health monitoring, while Qin Yimin et al. (2023) applied deep learning models to EEG data to predict cognitive workload.

In summary, head-mounted EEG devices provide valuable insights into brain activity, making them suitable for studying real-time emotional responses. However, most research has focused on clinical applications rather than media consumption. This study utilizes EEG technology to assess the effects of different podcast sounds on listener emotions, contributing to a deeper understanding of auditory-driven emotional responses. In this study, the FLEX 2 Saline - 32-channel wireless EEG headset system was used, which follows the international 10-20 electrode placement standard to ensure standardized brain activity measurements. Its wireless, saline design ensures high signal

fidelity and participant comfort, enabling accurate real-time EEG recordings of emotional responses generated by podcast voice modulation.

2.9.5 Indicators for Mood Monitoring

Mood monitoring relies on various indicators, including self-reported data, physiological signals, and machine learning-based emotion recognition systems. Cousins, Jon, and Michael (2020) developed a sentiment assessment method using predefined mood indicator descriptors, while Richard et al. (2019) introduced a mood-scoring system that integrates multiple emotional indicators for comprehensive evaluation.

Liu Zijun et al. (2019) explored mood monitoring devices that use sensors and AI-based analysis to detect real-time emotional states. Similarly, Rebecca Lietz et al. (2019) examined multimodal sentiment detection, incorporating speech, facial expressions, and physiological responses. Jonathon E. Giftakis et al. (2022) investigated neural indicators of mood by analyzing EEG-based emotion circuits.

In summary, mood monitoring technologies have advanced significantly, offering a combination of self-reported and objective physiological measures. While current research provides valuable frameworks for assessing emotions, there is limited exploration of mood tracking in podcast listening.

2.9.6 EEG Correlates of the 7 Positive Emotions

The study of 7 positive emotions—joy, excitement, contentment, love, hope, gratitude, and passion—has gained increasing attention in affective neuroscience, with EEG providing valuable insights into their neural correlates. Klimesch et al. (2020) found that theta (4–8 Hz) and beta (12–30 Hz) oscillations in the frontal and temporal regions are associated with joy and excitement, reflecting heightened cognitive engagement and reward processing. Similarly, León et al. (2021) demonstrated that love and gratitude, which involve social bonding and emotional connection, are linked to increased alpha (8–12 Hz) synchronization in the prefrontal cortex, indicative of emotional regulation. Furthermore, Moll et al. (2019) reported that gratitude is associated with delta (1–4 Hz) activity in the medial prefrontal cortex, which plays a crucial role in long-term emotional well-being and prosocial behavior.

Hope and passion, both future-oriented and motivational emotions, show distinct neural patterns. Lutz et al. (2020) observed that high-frequency gamma (30–45

Hz) activity in the prefrontal cortex is strongly correlated with goal-directed behavior and cognitive flexibility, supporting the role of hope in resilience and optimism. Demos et al. (2021) further found that passion, particularly in performance-based activities, is linked to sustained beta and gamma synchronization in the parietal and occipital lobes, suggesting an interplay between cognitive effort and mental imagery. In contrast, contentment is characterized by alpha power increases in the occipital and parietal lobes, reflecting a state of relaxation and internalized focus, distinguishing it from high-arousal emotions (Klimesch et al., 2020).

In summary, EEG research has established distinct brainwave signatures for the 7 positive emotions, with theta and beta waves dominant in joy and excitement, alpha and delta oscillations in love and gratitude, and gamma activity in hope and passion. Contentment, as a low-arousal emotion, is primarily associated with alpha synchronization in sensory regions.

2.10 Conceptual Framework for Research

Conceptual Framework for "Audio Voice Model to Reflect Audience Positive Emotion for Podcasting" The conceptual framework proposes that the vocal characteristics of podcast hosts, manipulated by the Audio Voice Model, and the gender characteristics of these hosts serve as independent variables. These variables influence the audience's positive emotional responses, which are measured as the dependent variable through self-report surveys and EEG indicators.

2.10.1 Independent Variables

1) Host Vocal Characteristics: This includes elements such as pitch, volume, speech rate, rhythm, timbre, accent, and pauses. The Audio Voice Model adjusts these parameters to produce a voice that reflects positive emotions.

2) Host Gender: Male and female hosts may exhibit different vocal traits that affect how their emotional expression is perceived.

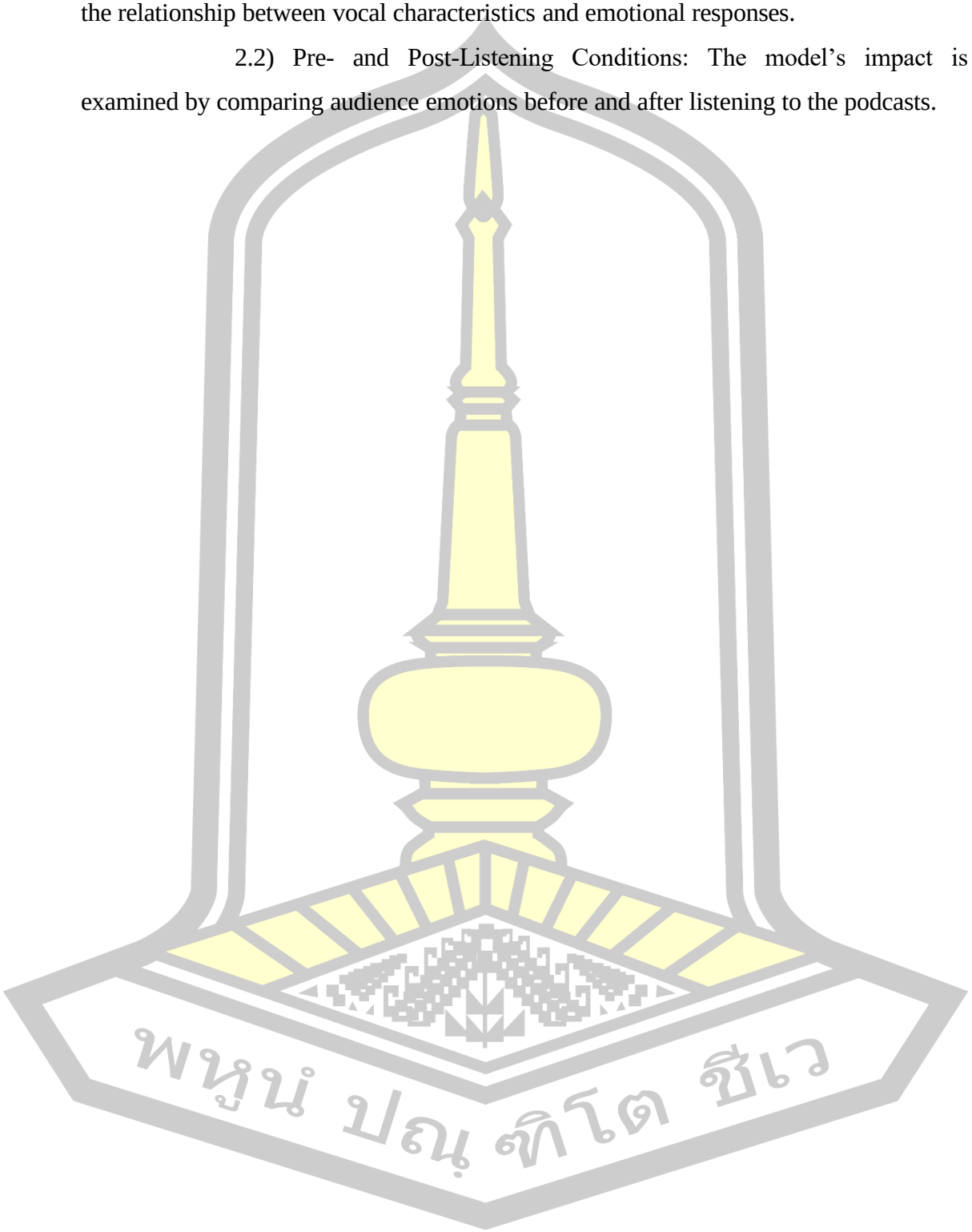
2.10.2 Dependent Variable

1) Audience Positive Emotional Responses: The degree to which listeners experience positive emotions such as joy, excitement, contentment, and gratitude. These responses are assessed through subjective evaluations and EEG data.

2) Moderating/Contextual Factors:

2.1) Listener Demographics: Age and listening frequency may moderate the relationship between vocal characteristics and emotional responses.

2.2) Pre- and Post-Listening Conditions: The model's impact is examined by comparing audience emotions before and after listening to the podcasts.



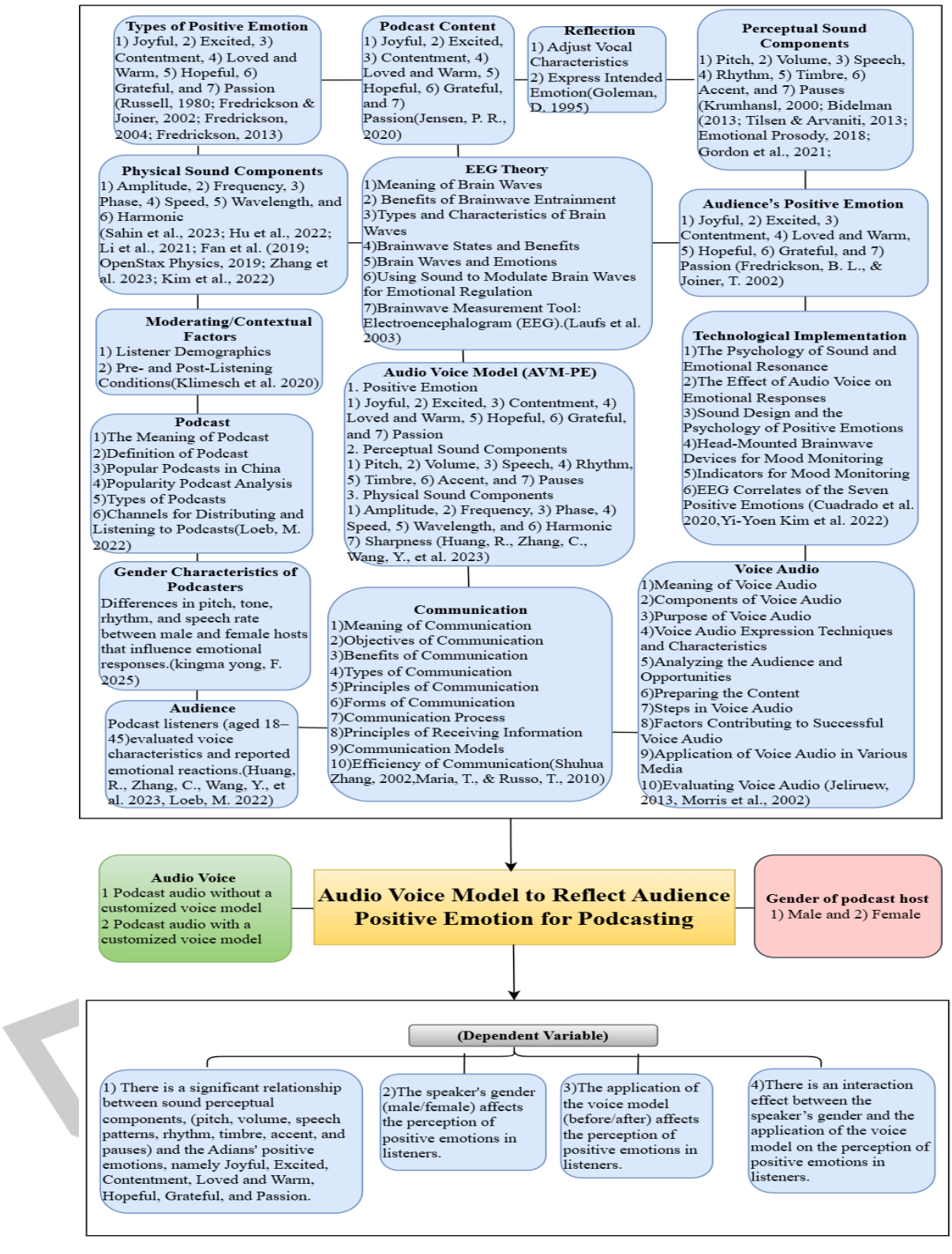


Figure 2. 15 The Conceptal Framework of Researche

CHAPTER III

Research Methodology

This research method used a mixed method approach, including surveys and expert interviews, and covered adult male and female audiences of podcast program, the most popular positive podcasts with male and female hosts, and the quality of the model. The following research steps were implemented.

3.1 Research Objectives

3.1.1 To investigate the impact of the voices and genders of podcast hosts on the audience's positive emotional responses.

3.1.2 To develop an audio voice model to reflect the audience's positive emotions for podcasting

3.1.3 To investigate the correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts.

3.1.4 To compare of the audience’s positive emotions before and after applying the voice model to podcast hosts of different genders.

3.2 Research Hypothesis

3.2.1 There is a significant relationship between sound perceptual components, (pitch, volume, speech patterns, rhythm, timbre, accent, and pauses) and the Adians' positive emotions, namely Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.

3.2.2 The speaker's gender (male/female) affects the perception of positive emotions in listeners.

3.2.3 The application of the voice model (before/after) affects the perception of positive emotions in listeners.

3.2.4 There is an interaction effect between the speaker’s gender and the application of the voice model on the perception of positive emotions in listeners.

3.3 Population and Samples

3.3.1 Population

The target population for this research consists of 2 main groups:

1) Podcast Listeners – Individuals who regularly engage with podcast content, representing a diverse audience in terms of age, gender, and listening preferences.

2) Podcast Hosts – Professional or amateur podcasters who produce and present content, covering a range of speaking styles and vocal characteristics.

3.3.2 Sample

To ensure a representative sample, the study will include:

1) 50 podcast listeners who actively engage with podcast content at least once a week.

2) 10 podcast hosts (5 male and 5 female) who have experience in hosting podcasts and can provide varied vocal characteristics for analysis.

The sample size is determined based on previous studies in voice and emotion research, ensuring statistical reliability while maintaining feasibility in data collection and EEG analysis.

3.3.3 Sampling Method

1) For podcast listeners, a stratified random sampling method will be applied to balance demographic factors such as age and gender, ensuring the sample is representative of the general podcast audience.

2) A purposive sampling technique will be used for selecting podcast hosts to ensure diversity in vocal characteristics (e.g., gender, pitch, speech rate, timbre).

This approach enhances the validity of the study by selecting participants who are relevant to the research objectives while minimizing selection bias.

3.4 Risk Reduction for Volunteers

This research aims to collect data from two groups of volunteers: podcast listeners and podcast hosts. The researcher prioritizes risk prevention for volunteers before, during, and after their participation in the study. To ensure their safety, the following inclusion and exclusion criteria have been established:

3.4.1 Volunteer Inclusion Criteria

1) Podcast Audience

1.1) Aged between 18–45 years

1.2) Listens to podcasts at least 1–3 times per week

1.3) Has the ability to comprehend podcast content in Chinese

1.4) Has no hearing impairments or neurological conditions affecting sound perception.

1.5) Willing to participate in the study and provide necessary data

2) Podcast Hosts

2.1) Has at least 1 year of experience as a podcast host

2.2) Capable of adjusting vocal characteristics based on the Audio Voice Model

2.3) Proficient in voice modulation, including pitch, volume, and rhythm control

2.4) Willing to participate and follow the designated voice design guidelines

3.4.2 Exclusion Criteria

1) Podcast Audience

1.1) Has hearing impairments or neurological disorders

1.2) Cannot comprehend Mandarin-language podcast content

1.3) Has a history of severe emotional disorders affecting mood measurement

2) Podcast Hosts

2.1) Lacks experience as a podcast host

2.2) Unable to modify voice elements according to study requirements

2.3) Cannot consistently participate throughout the research process

3.5 Research Variable

3.5.1 Independent Variables

The independent variables in this study are the factors that influence the audience's emotional responses:

1) Podcast Hosts' Vocal Characteristics include: Pitch, Volume, Speech Rate, Rhythm, Timbre, Accent, and Pauses

2) Gender of Podcast Hosts (Male/Female)

3) Audio Voice Model (AVM-PE) Implementation has 6 components: 1) Input Audio Processing, 2) Audio Analysis, 3) Emotion Adaptive Voice Synthesis, 4) Creative Voice Modulation, 5) Output Audio Generation, 6) Evaluation and Optimization

3.5.2 Dependent Variables

The dependent variables are the measurable outcomes that reflect the impact of the independent variables:

1) The impact of the voices and genders of podcast hosts on the audience's positive emotional responses. Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.

2) The Suitability of the audio voice model to reflect the audience's positive emotions for podcasting

3) The correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts.

4) The results of the compare of the audience's positive emotions before and after applying the voice model to podcast hosts of different genders.

3.6 Research Tools

3.6.1 Research Tools

1) Podcast Impact Evaluation: Evaluated the effect of the voice and gender of podcast hosts on positive emotional responses.

2) Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

3) Evaluation Form: Assessed the suitability of the components of an audio voice model for reflecting audience positive emotions in podcasts.

4) EEG Assessment: Measured the audience's positive emotions before and after listening to podcasts using Electroencephalography (EEG).

5) Semi-Structured Interview: Evaluated the relationship between the audio voice model and podcast host gender on positive emotions.

3.6.2 Research Equipment

1) Voice Analysis Software: analysed the components of podcast audio.

2) The FLEX 2 Saline - 32 Channel Wireless EEG Head Cap System is a head-mounted brainwave meter designed to capture EEG signals in real time, enabling the measurement of listeners' positive emotional responses to different podcast programs.

Utilizing the international 10-20 electrode placement system, this system ensures standardized and reliable brain activity recording.

Its 32-channel configuration provides comprehensive brain coverage, capturing neural activity relevant to auditory and emotional processing. By integrating EEG data with the Positive Emotions Measurement Scale, this study conducted a detailed assessment of listeners' emotional experiences, analyzing pleasure, excitement, and other affective responses to podcast content.

Figure 3. 1 FLEX 2 Saline - 32 Channel Wireless EEG Head Cap System

3) Voice Recording and Processing Software: Tools for capturing and analyzing voice data to develop the model.

3.7 Tools Development and Quality Assessment

3.7.1 Develop and Verify the Quality of Research Tools

1) Podcast Impact Evaluation: The Effect of Voice and Gender of Podcast Hosts on Positive Emotional Responses

Survey questionnaire form: This questionnaire aimed to evaluate participants' emotional responses to the voice and gender of podcast hosts. By evaluating different aspects of voice, such as pitch, volume, rhythm, timbre, pauses, speech rate, and accent, as well as the gender of the podcast host, this study sought to measure the emotional impact that these factors had on listeners' perceptions and experiences.

The questionnaire was divided into four sections. Section 1 collected demographic information. Section 2 assessed the impact of voice characteristics on

participants' emotions. Section 3 examined how the host's gender influenced participants' emotions. Section 4 provided space for optional comments.

The survey consisted of four sections:

Section 1: Basic information, which included a 4-item checklist (gender, age, podcast listening frequency per week, and preferred gender of the podcast host).

Section 2: Positive Emotional Response to the Host's Voice Characteristics, which required participants to rate how each voice characteristic of the podcast host evoked specific positive emotions, using a 5-point scale.

Table 3. 1 Sound characteristics Evaluation criteria for positive emotional responses in listeners

Evaluation criteria for the degree of motivation	**Interpretation Criteria**
5 = Extremely	4.51-5.00 = Extremely
4 = Very much	3.51-4.50 = Very much
3 = Moderately	2.51-3.50 = Moderately
2 = Slightly	1.51-2.50 = Slightly
1 = Not at all	1.00-1.50 = Not at all

Section 3: Impact of the Host's Gender on Emotional Responses

The gender of the podcast host influenced the positive emotional responses of the interviewees to varying extents. Participants rated their responses using a 5-point scale.

Table 3. 2 Assessment criteria for the influence of the moderator's gender on the emotional response of the audience

Impact Assessment Table	**Interpretation Criteria**
5 = Strongly agree	4.51-5.00 = Strongly agree
4 = Agree	3.51-4.50 = V Agree
3 = Neutral	2.51-3.50 = Neutral
2 = Disagree	1.51-2.50 = Disagree
1 = Strongly disagree	1.00-1.50 = Strongly disagree

Section 4: Additional Comments This section allowed respondents to provide suggestions beyond the evaluated items.

Audio Feature Extraction and Measurement: Adobe Audition and Praat were used for audio analysis to extract the 7 core speech features that influence listener emotions, and a unified measurement standard was developed to ensure the scientific and comparable nature of the data.

In this study, 10 podcast hosts were selected as the source of audio samples, including 5 females (F1-F5) and 5 males (M1-M5). The audio material was taken from the actual released podcasts, program standard: sampling rate of 44.1 kHz, 16-bit stereo format to ensure consistency of sound quality, audio duration: the length of each segment was no more than 3 minutes, with a total of 10 segments.

Audio Feature Extraction**:** Adobe Audition and Praat were used for audio analysis and feature extraction, evaluating the following 7 core vocal characteristics based on numerical evaluation standards:

Table 3. 3 Audio Feature Extraction and Standards for Broadcast Programs

Audio Feature	**Computation Method**	**Evaluation Standard**	**Reference Range**	**Interpretation**
Amplitude	Peak signal intensity calculation	High: 0.8-1.0, Medium: 0.6-0.8, Low: <0.6	0.5 - 1.0	Determines the loudness of the audio; higher amplitude indicates stronger energy, while lower amplitude results in softer audio.
Frequency	Mean fundamental frequency calculation	High: >300 Hz, Medium: 150-300 Hz, Low: <150 Hz	50 - 350 Hz	Affects pitch; higher frequencies are linked to excitement, while lower frequencies are associated with stability.
Phase	Initial waveform phase measurement	Range: -3.14 to 3.14	-3.14 - 3.14	Affects clarity and resonance of the sound; stable phase provides clearer and more natural audio.

Audio Feature	Computation Method	Evaluation Standard	Reference Range	Interpretation
Speed	Words per minute (WPM) calculation	Fast: >160 BPM, Medium: 100-160 BPM, Slow: <100 BPM	80 - 200 BPM	Faster speech rates convey dynamism and urgency, while slower speech rates create a calm atmosphere.
Wavelength	Primary cycle length measurement	Short: <200 ms, Medium: 200-600 ms, Long: >600 ms	100 - 1000 msShort wavelengths correspond to sharper and more energetic voices, while longer wavelengths produce deeper and steadier tones.	Short wavelengths correspond to sharper and more energetic voices, while longer wavelengths produce deeper and steadier tones.
Harmonic	Spectral centroid frequency calculation	Bright: >150 Hz, Neutral: 100-150 Hz, Warm: <100 Hz	50 - 200 Hz	Higher harmonic frequencies result in clearer and more resonant audio, while lower harmonic frequencies create a warmer and richer tone.

Audio Feature	Computation Method	Evaluation Standard	Reference Range	Interpretation
Sharpness	Sharpness index calculation	Sharp: >1.5, Neutral: 1.0-1.5, Smooth: <1.0	0.1 - 2.0	Higher sharpness makes the voice more piercing and distinct, whereas lower sharpness results in a smoother and softer sound.

Different podcast audio frequencies affect listeners' brainwave activity and emotional states. First, the EEG (electroencephalogram) measurement device was used to record the brainwave changes of listeners under different audio stimuli, focusing on the response characteristics of the Alpha (8-13 Hz), Beta (13-30 Hz), Theta (4-8 Hz) and Gamma (30+ Hz) frequency bands. Second, the audio frequency analysis tool was used to extract features of the podcast audio signal, including core parameters such as amplitude variation, waveform structure and oscillation mode. In addition, in order to ensure the accuracy of data acquisition, a synchronization control system is used to strictly synchronize the audio playback with the EEG recording to avoid external interference. The combination of this series of validation tools enables the study to accurately capture the physiological responses of listeners to different audio frequencies, providing a reliable basis for subsequent data analysis.

2) Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

Steps for Developing an Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

Step 1: Define Objectives and Parameters

1 . Identified the primary goal of the model: to enhance positive emotional responses in podcast audiences.

2 . Determined key parameters, including tone, pitch, pace, clarity, and emotional resonance of the voice.

Step 2: Review Literature and Best Practices

1. Conducted a literature review on voice modulation, emotional perception, and podcasting techniques.

2. analysed successful case studies of voice applications in media and communication.

Step 3: Analyze Audience Preferences

1. Conducted surveys and focus groups to gather audience preferences regarding voice qualities that evoke positive emotions.

2. Collected data on gender-specific preferences and emotional triggers.

Step 4: Recruit and Train Podcast Hosts

1. Selected 10 podcast hosts (5 male, 5 female) with at least 3 years of experience in content delivery.

2. Provided training on voice modulation techniques, focusing on achieving consistency in tone, pitch, and emotional engagement.

Step 5: Develop the Voice Model Framework

1. Designed a framework for the Audio Voice Model, incorporating insights from research and audience data.

2. Included modules for tone adjustments, emotional inflections, and dynamic delivery.

Step 6: Implement Voice Analysis Tools

1. Integrated tools for real-time voice analysis to measure tone, pitch, and emotional delivery.

2. Used software to monitor and provide feedback on the hosts' performance during training sessions.

Step 7: Test the Model with Pilot Groups

1. Conducted pilot tests with small audience groups to evaluate the model's effectiveness.

2. Collected feedback on the emotional impact and overall listening experience.

Step 8: Refine the Model

1. analysed data from pilot tests and adjusted the model parameters as needed.

2. Iterated the training process for hosts to align their delivery with the refined model.

Step 9: Final Validation

1. Conducted a larger-scale test with 50 audience participants to validate the model's performance.

2. Measured positive emotional responses using EEG tools and post-listening surveys.

Step 10: Documentation and Deployment

1. Documented the development process and finalized the Audio Voice Model guidelines.

2. Deployed the model for use in podcast production, ensuring consistency and scalability.

3) Evaluation Form for the Suitability of the Components of an Audio Voice Model for Reflecting Audience Positive Emotions in Podcasts

Quality Format Assessment Survey: This questionnaire selected 5 experts (including professionals who provided insights from both technical and psychological perspectives) to evaluate the appropriateness of the components of the audio voice model that reflected the positive emotions of the audience in the podcast.

3.1） Audio engineer/designer

Expertise: sound design, voice modulation, and how audio affects emotions.

3.2） Media psychologist

Expertise: emotional and psychological effects of media content, especially speech and sound.

3.3） Voice coach/actor

Specialty: voice modulation, pitch control and the emotional impact of voice delivery.

3.4） Podcast producer

Specialty: creating emotionally engaging audio content and ensuring that the voice resonates with the podcast audience.

3.5） Communication scholar

Specialty: researching sound characteristics and their role in emotional communication.

The questionnaire used a mixed-method approach, collecting both quantitative and qualitative data. It was divided into the following three sections:

Section 1: Basic information about the respondent, which included a 6-item checklist (gender, age, highest educational qualification, field of study, current occupation, and years of experience in audio design, podcasting, or related fields).

Section 2: A series of questions that evaluated the appropriateness of 10 components of the audio voice model for reflecting audience positive emotions in podcasts.

Table 3. 4 Assessment criteria for the suitability of audio sound model components

Evaluation Criteria for Opinions	**Interpretation Criteria for Opinions**
5 = Strongly agree	4.51-5.00 = Strongly agree
4 = Agree	3.51-4.50 = V Agree
3 = Neutral	2.51-3.50 = Neutral
2 = Disagree	1.51-2.50 = Disagree
1 = Strongly disagree	1.00-1.50 = Strongly disagree

Section 3: Additional suggestions and feedback

3.6) The experts' responses served as a validation of the model's appropriateness and provided certification for the Audio Voice Model.

3.7) The completed questionnaire was used to improve and refine the audio speech model to ensure its completeness and applicability.

4) Assessment of Audience's Positive emotions Before and After Listening to Podcasts Using Electroencephalography (EEG)

Positive Emotion Measurement Form : This head-mounted device was an advanced brainwave capture tool that measured the emotional responses of listeners in real time before and after listening to podcasts. By recording brain activity, the study provided a more intuitive understanding of the degree of listeners' emotional responses to different programmes. At the same time, participants completed a positive emotion measurement scale, which enabled a more comprehensive evaluation of their emotional experiences. This approach allowed the research to explore in greater depth the impact of podcast programmes on listeners' positive emotions.

4.1) Data were recorded before and after testing each listener.

4.2) Researchers selected 10 popular and positive content pieces from the Himalaya platform, with a total of 10 episodes. The listening time for each episode was controlled to less than 3 minutes.

Table 3. 5 List of podcasts with positive content recommended by male and female hosts

Top 5 Podcasts Hosted by Males	**Top 5 Podcasts Hosted by Females**
M1. Positive Energy Station - Short, uplifting episodes sharing motivational insights.	F1. The Optimism Channel - Bite-sized sessions focusing on positive thinking.
M 2. Mind Boost Radio - Quick tips on mental wellness and positivity, perfect for a brief listen.	F 2. Sunshine Stories - Heartwarming tales delivered in under 3 minutes.
M 3. Uplifting Moments - Daily 3-minute inspiration to start your day positively.	F 3. Feel Good Talk - Quick, practical tips for living a happier life.
M 4. Better Days Ahead - Concise episodes with advice on staying positive.	F 4. Positive Mind Vibes - Short sessions on well-being and positive energy.
M 5. Happiness Radio - Joyful snippets to brighten up your day.	F 5. Hopeful Hearts - Brief, encouraging stories of hope and resilience.

Section 1: Basic information, using a 4-item checklist (gender, age, Highest Educational Qualification, Occupation).

Section 2: Male and Female Hosts' Assessment of Positive Emotions Before and After Listening to Podcasts 7 positive emotions were assessed using electroencephalography (EEG) to monitor real-time changes in positive emotions while listeners listened to 10 different podcasts. EEG recorded the frequency of the listeners' brainwaves and collected data on their emotional states.

Table 3. 6 (EEG) frequency bands and functional interpretation criteria for emotional responses

Frequency Band	**Frequency Range (Hz)**	**Power Value Range (a.u.)**	**Functional interpretation standard**
Delta (δ)	0.5–4 Hz	4–6 a.u.	High power indicates enhanced deep emotional processing (e.g., joy, excitement), reflecting active unconscious emotional memory or processing.
		2–3 a.u.	Low power suggests weaker emotional arousal and reduced relaxation.

Frequency Band	Frequency Range (Hz)	Power Value Range (a.u.)	Functional interpretation standard
Theta (θ)	4–8 Hz	3–5 a.u.	High power reflects significant memory consolidation and emotional regulation, often seen in highly engaging emotional experiences like joy or creative thinking.
		1.5–2.5 a.u.	Low power indicates less active emotional regulation or memory processing.
Alpha (α)	8–12 Hz	1–2 a.u.	Power below 1 a.u. indicates increased positive emotional engagement, reducing relaxation; stable power suggests lower emotional engagement with a relaxed and alert state.
Beta (β)	12–30 Hz	3–5 a.u.	High power shows increased cognitive engagement, typically seen during excitement or task-related emotional states; low power suggests reduced cognitive load or emotional fluctuation.
		1–3 a.u.	Indicates normal cognitive task states with minimal emotional or task-related impact.
Gamma (γ)	30–45 Hz	>2 a.u.	Significantly high power reflects active complex emotional or higher-order cognitive processing, such as passion or deep emotional content engagement.
		0.5–2 a.u.	Low power indicates limited emotional or cognitive engagement, with less demand for complex information processing.

5) Semi-Structured Interview for Evaluating the Correlations Between Audio Voice Model and Podcast Host Gender on Positive

Emotions Interview Form Through semi-structured interviews, researchers gained an in-depth understanding of listeners' responses to positive emotions and their feelings about the vocal characteristics of announcers of different genders. The interviews provided deeper insights into listeners' perspectives and experiences and explored the relationship between audio-speech models and listeners' gender

characteristics. The focus was on how these characteristics affected the effectiveness of audio-speech models in reflecting positive emotions.

Data collection involved interviews with participants who provided detailed insights into their emotional responses and perceptions of voice characteristics. The following steps were taken to develop the interview guide:

5.1) Reviewing Concepts and Principles:

The researcher reviewed key concepts and principles related to audio voice models, vocal characteristics, and their emotional impact on listeners to establish the scope and content of the interview guide.

5.2) Defining Interview Topics: The interview topics were divided into two sections:

Section 1: Basic information, using a 4-item checklist (gender, age, Highest Educational Qualification, Occupation).

Section 2: Structured questions focused on participants' perceptions of:

1) Male and female podcast hosts' vocal characteristics;

2) Emotional responses to different voice characteristics;

3) Alignment between voice characteristics and positive emotional experiences;

4) The role of audio voice models in enhancing emotional engagement;

5) Potential improvements in podcasting voice design.

5.3) Quality Assurance: The interview guide was reviewed and revised by the researcher to ensure alignment with the study's objectives, clarity, and coherence. Feedback from preliminary pilot interviews was used to refine the questions.

5.4) Pilot Testing: The guide was tested with a small group of participants to validate its effectiveness in eliciting relevant responses and ensuring alignment with the research objectives.

5.5) Final Adjustments: Based on the pilot testing feedback, the interview guide was finalized and prepared for implementation with the main participant group.

3.8 Risk Mitigation for Volunteers

This study received ethical approval under the code 599-510/2024. Its purpose was to collect data from volunteers interested in listening to online podcasts on the

Himalaya Podcast platform. To minimize risks that might arise for those involved in this research, the researchers established the following inclusion and exclusion criteria:

3.8.1 Inclusion Criteria

1) Volunteers must be at least 18 years old.

2) Participants must show interest in listening to positive-themed podcasts.

3) Volunteers must demonstrate proficiency in understanding spoken Chinese.

4) Participants must willingly provide informed consent to join the research voluntarily.

5) Volunteers must be in good physical and mental health.

3.8.2 Exclusion Criteria

Volunteers were excluded if they missed participation in two consecutive activities.

3.9 Experimental Action Plan

Experimental Grouping by the researcher planned the experimental grouping as follows:

3.9.1 Preparation of Podcasts

1) Selected 5 male and 5 female podcast hosts, totaling 10 participants, based on the eligibility criteria for volunteer selection.

2) Prepared 7 stories, each lasting 3 minutes, designed to reflect 7 positive emotions. The 10 podcast hosts delivered these stories in 2 formats:

2.1) Format A – The 10 podcast hosts narrated the 7 positive emotion stories before applying the audio voice model.

2.2) Format B – The same 10 podcast hosts narrated the 7 stories using the audio voice model for voice enhancement.

3) Recorded the voices of the ten podcast hosts as they narrated the 7 positive emotion stories before applying the audio voice model.

4) Trained the 10 podcast hosts to use the audio voice model to reflect the audience's positive emotions in podcasting.

5) Recorded the voices of the ten podcast hosts narrating the same 7 stories using the audio voice model for voice enhancement.

6) Selected 5 0 volunteers as podcast listeners based on the specified criteria, ensuring voluntary participation.

3.9.2 Experimental Procedures

Table 3. 7 Detailed Procedures for Experimental Phases

Phase		**Step**	**Details**	**Time**	**Notes**
1. Pre-Listening Phase	1.1	Explain Experiment	Researcher explained objectives, answered questions, and obtained consent from 50 volunteers.	10-15 min	Ensured participants fully understood before proceeding.
	1.2	Preliminary Questionnaire	Volunteers completed a questionnaire to assess emotions before listening.	5-10 min	Measured baseline positive emotions.
	1.3	Group Division	Participants divided into two groups: Group M (5 male hosts) and Group F (5 female hosts).	5 min	Ensured balanced experimental groups.
	1.4	Audience Composition	Total 50 participants: 25 males and 25 females.	N/A	Demographic balance maintained.
	2.1	Segment 1 - Male Hosts	Listeners heard 7 positive emotion stories narrated by male hosts under 2 formats.	105 min	Each listener experienced both formats.
		2.1.1 Format A - Male Hosts	Male hosts narrated stories before applying an audio voice model (3 min each, total 105 min).	105 min	Control condition without enhancement.
		2.1.2 Format B - Male Hosts	Male hosts narrated stories using an enhanced audio voice model (3 min each, total 105 min).	105 min	Experimental condition with enhancement.
		2.1.34EEG Data Collection (Male)	EEG device recorded emotional responses of listeners.	During Listening	Real-time emotional data collected.
		2.1.3 Questionnaire	Participants completed a	5-10 min	Measured post-

Phase		Step	Details	Time	Notes
		(Male)	post-listening questionnaire (5-10 min).		listening emotions.
2. Listening Phase	2.2	Segment 2 - Female Hosts	Listeners heard 7 positive emotion stories narrated by female hosts under 2 formats.	105 min	Each listener experienced both formats.
		2.2.1 Format A - Female Hosts	Female hosts narrated stories before applying an audio voice model (3 min each, total 105 min).	105 min	Control condition without enhancement.
		2.2.2 Format B - Female Hosts	Female hosts narrated stories using an enhanced audio voice model (3 min each, total 105 min).	105 min	Experimental condition with enhancement.
		2.2.3 EEG Data Collection (Female)	EEG device recorded emotional responses of listeners.	During Listening	Real-time emotional data collected.
		2.2.4 Questionnaire (Female)	Participants completed a post-listening questionnaire (5-10 min).	5-10 min	Measured post-listening emotions.
	3.1	EEG Data Analysis	EEG data analysed to compare emotions before and after voice enhancement.	Varied	Key comparison for voice effect.
	3.2	Sound Component Analysis	analysed relationship between sound components and positive emotions.	Varied	Investigated how sound influences emotions.
3. Post-Listening Phase	3.3	Gender Effect Analysis	analysed effect of podcast host gender on emotional responses.	Varied	Identified gender-based emotional differences.
	3.4	Response Difference Analysis	Compared emotional responses before and after using voice enhancement.	Varied	Compared pre- and post-enhancement responses.
	3.5	Interaction Analysis	Examined interaction between host gender and voice enhancement.	Varied	Tested combined effects of gender and voice.

Phase	Step	Details	Time	Notes
	3.6 Result Compilation	Summarized experimental results and compiled research report.	Varied	Finalized study findings.

From Table 3.7, the overall experimental phase is shown and the information below is the details of the experimental phase.

1) Pre-Listening Phase

1.1) The researcher explained the objectives of the experiment to the 50 volunteers who participated in the study and asked additional questions to ensure understanding before obtaining their consent for participation from the beginning to the end of the experiment.

1.2) The volunteers completed a preliminary questionnaire to assess their positive emotions before listening to two types of podcasts.

1.3) The experiment groups were divided into two groups:

1.3.1) Group M: Podcasts hosted by 5 male hosts

1.3.2) Group F: Podcasts hosted by 5 female hosts

1.4) The audience consisted of 50 listeners, including 25 males and 25 females.

2) Listening Phase

2.1) Segment 1: The 50 listeners listened to podcasts presented by 5 male hosts, who narrated 7 positive emotion stories under 2 formats:

2.1.1) Format A – The male hosts narrated the 7 positive emotion stories before applying the audio voice model for 3 minutes each, totaling 105 minutes.

2.1.2) Format B – The male hosts narrated the same 7 stories using the audio voice model for voice enhancement for 3 minutes each, totaling 105 minutes.

2.1.3) During this phase, the researcher collected emotional response data using an EEG device.

2.1.4) The 50 participants completed a questionnaire to assess their positive emotions after listening to the 2 formats of podcasts for 5-10 minutes.

2.2) Segment 2: The 50 listeners listened to podcasts presented by 5 female hosts, who narrated 7 positive emotion stories under 2 formats:

2.2.1) Format A – The female hosts narrated the 7 positive emotion stories before applying the audio voice model for 3 minutes each, totaling 105 minutes.

2.2.2) Format B – The female hosts narrated the same 7 stories using the audio voice model for voice enhancement for 3 minutes each, totaling 105 minutes.

2.2.3) During this phase, the researcher also collected EEG data to monitor the emotional responses of the listeners.

2.2.4) The 50 participants completed a questionnaire to assess their positive emotions after listening to the 2 formats of podcasts for 5-10 minutes.

3) Post-Listening Phase

3.1) The researcher analysed the EEG data to compare the levels of positive emotion from the voice before and after applying the enhancement audio voice model.

3.2) The researcher analysed the relationship between perceptual sound components (pitch, volume, speech, rhythm, timbre, accent, and pauses.) and positive emotions, including joyful, excited, contentment, loved and warm, hopeful, grateful, and passion.

3.3) The researcher analysed the effect of the podcast host's gender (male/female) on the audience’s positive emotional responses.

3.4) The researcher analysed the differences in the audience’s positive emotional responses before and after using the voice model.

3.5) The researcher analysed the interaction between the podcast host's gender and the use of the voice model on the audience’s positive emotional responses.

3.6) Summarize the experimental results and compile a report for research dissemination.

3.10 Data Collection

The data collection process for this research was conducted in four phases as follows:

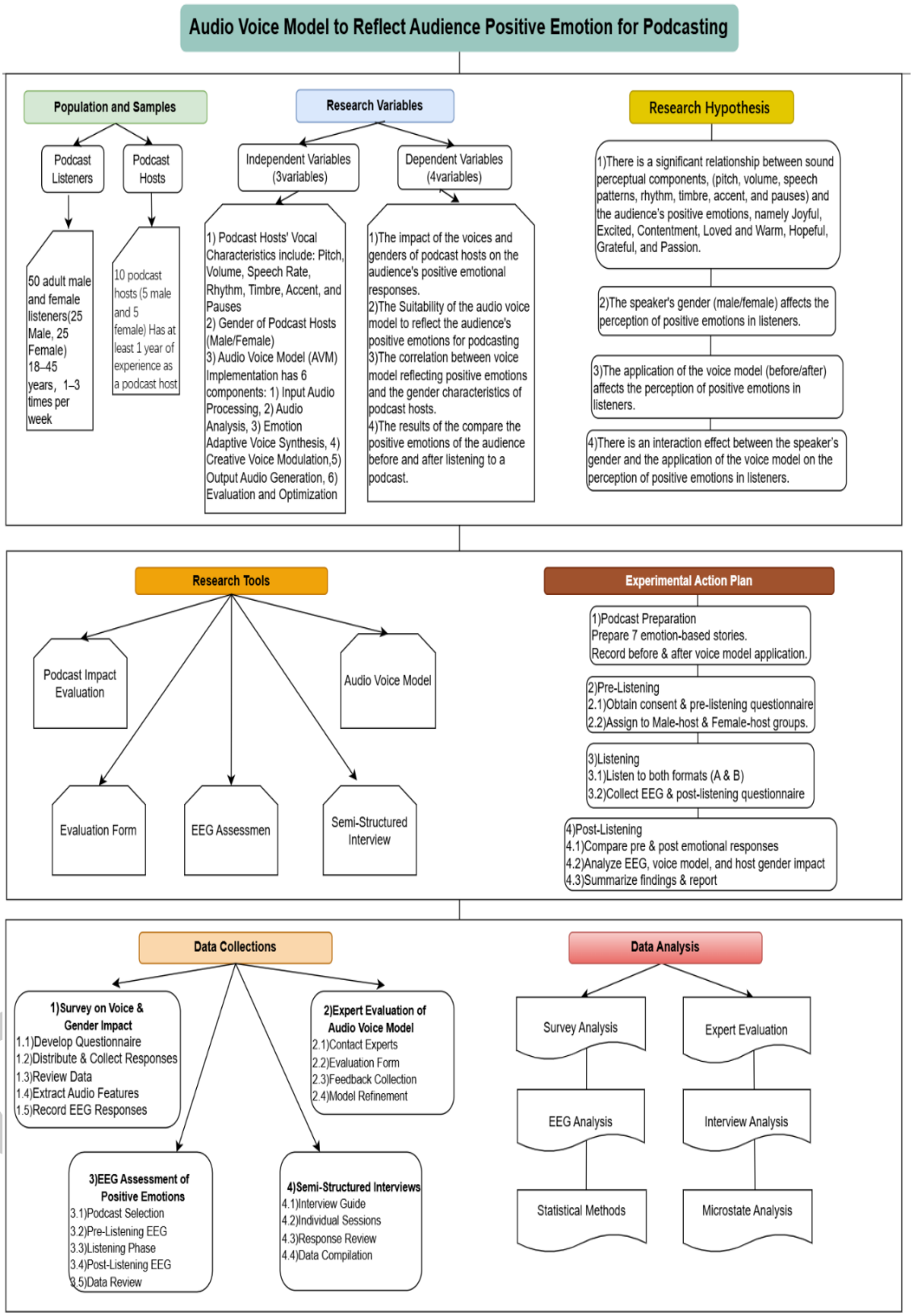


Figure 3. 2 Data Collection Process Chart

3.10.1 Survey on the Impact of Voice and Gender of Podcast Hosts on Positive Emotional Responses

Data were collected from participants using a survey questionnaire designed to evaluate their emotional responses to the voice and gender of podcast hosts. The data collection procedure was as follows:

1) The researcher developed the survey questionnaire, which included four sections: demographic information, emotional responses to voice characteristics, gender influence on emotions, and additional comments.

2) The questionnaire was distributed online through popular social media platforms to reach the target audience.

3) Participants provided voluntary consent before completing the survey.

4) After achieving the target number of responses, the researcher reviewed the data for completeness and accuracy before proceeding with analysis.

5) Extract audio features using Adobe Audition and Prat to analyze 7 core speech characteristics that influence listener mood. The features extracted were Amplitude, Frequency, Phase, Speed, Wavelength, Harmonic, and Sharpness. Each feature is calculated using a fixed method to ensure consistency and comparability of all audio data.

6) An EEG (electroencephalography) device was used to record the brainwave activity of the subjects while listening to podcasts at different audio frequencies. The experimental environment was kept quiet, and the subjects wore the EEG device and listened to different frequencies of podcasts, including 187.63 Hz, 264.55 Hz, 169.45 Hz, etc., under the same conditions, and the playback time of each audio segment was fixed. During the audio playback process, the EEG device recorded the subjects' EEG signals in real time and stored the data automatically to ensure data integrity and consistency. All audios were standardized to control the interference of non-audio variables.

3.10.2 Evaluation of the Suitability of Audio Voice Model Components

Data were collected from 5 experts to assess the suitability of audio voice model components in reflecting positive emotions in podcasts. The data collection steps were as follows:

1) The researcher contacted 5 experts from relevant fields (audio engineering, media psychology, voice coaching, podcast production, and communication studies) and scheduled a time for completing the evaluation form.

2) Experts were provided with the evaluation form, which included questions on ten audio voice model components and additional feedback sections.

3) Experts completed the evaluation form, and the researcher reviewed their responses for completeness.

4) The researcher used the feedback to refine the audio voice model and ensure its alignment with the study's objectives.

3.10.3 Assessment of Audience's Positive Emotions Before and After Listening to Podcasts

Data were collected using EEG to measure participants' positive emotional responses before and after listening to podcasts. The data collection steps were as follows:

1) The researcher prepared a list of ten preselected podcast episodes (5 hosted by males and 5 by females) and briefed participants on the study's procedure.

2) EEG recordings were taken to establish baseline emotional data before podcast listening.

3) Participants listened to the podcast episodes, with each episode lasting no more than three minutes.

4) Post-listening EEG recordings were conducted to capture changes in emotional responses.

5) The researcher reviewed the EEG data for completeness and accuracy, focusing on brainwave frequency bands (delta, theta, alpha, beta, and gamma) to analyze emotional changes.

3.10.4 Semi-Structured Interviews on Audio Voice Models and Host Gender

Semi-structured interviews were conducted to explore the relationship between audio voice models and podcast host gender on positive emotional responses. The data collection steps were as follows:

1) The researcher developed an interview guide based on key concepts related to voice characteristics and emotional impact.

2) Participants were interviewed individually, and the sessions were recorded with their consent. The interviews focused on participants' perceptions of male and female podcast hosts' voice characteristics and their impact on emotions.

3) After completing the interviews, the researcher reviewed the responses for completeness.

4) The interview data were analysed to identify patterns and insights, which were incorporated into the study's findings.

3.11 Data Analysis

Data Analysis Procedures for the Research on "Audio Voice Model to Reflect Audience Positive Emotion for Podcasting" followed a structured sequence aligned with the research instruments and expected outcomes. The process was as follows:

3.11.1 Analysis of Basic Information of the Sample Group

1) Descriptive Analysis of Population and Sample

The demographic characteristics of the sample (gender, age, podcast listening frequency, and preferences regarding podcast host gender) were analysed. Descriptive statistics (percentage, mean, and standard deviation) were used to summarize the demographic data.

3.11.2 Analysis of the Impact of Podcast Host Voice and Gender on Audience's Positive Emotional Responses

1) Analysis of the Audience's Positive Emotion Questionnaire Before and After Listening to Podcasts

Descriptive statistics were used to present percentages, mean scores, and standard deviations of positive emotion scores. Paired t-test was applied to compare the levels of positive emotions before and after listening to podcasts.

2) Analysis of the Differences in Emotional Responses Based on Podcast Host Gender

Two-way ANOVA was conducted to compare the effects of male and female podcast hosts on audience positive emotional responses.

3.11.3 Analysis of the Relationship Between Sound Components and Positive Emotions

1) Analysis of the Audio Voice Characteristics of Podcast Hosts

The 7 audio components (pitch, volume, speech rate, rhythm, timbre, accent, and pauses) were analysed.

2) Analysis of the Correlation Between Sound Perception Components and Positive Emotions

Pearson correlation analysis was used to examine the relationship between independent variables (sound components) and dependent variables (7 positive emotions: joyful, excited, contentment, loved and warm, hopeful, grateful, and passion).

3.11.4 EEG Analysis of Brain Responses to Podcast Listening

1) Comparison of Brainwave Differences Before and After Listening to Podcasts

Paired t-test was used to compare EEG measurements in theta, alpha, beta, and gamma bands before and after listening.

2) Analysis of the Correlation Between Podcast Host Voice Characteristics and EEG Responses

Pearson correlation analysis was conducted to examine the relationship between vocal characteristics of podcast hosts and changes in EEG data.

3) Analysis of Gender Differences in EEG Responses to Podcast Hosts

Two-way ANOVA was applied to compare the effects of male and female podcast hosts on brainwave activity.

3.11.5 Analysis of the Impact of Using the Audio Voice Model

1) Comparison of Positive Emotions Before and After Using the Audio Voice Model

Paired t-test was used to compare the impact of the Audio Voice Model on audience emotional responses.

2) Comparison of the Effects of the Audio Voice Model on Male and Female Podcast Hosts

Two-way ANOVA was applied to analyze the interaction effect between podcast host gender and the use of the Audio Voice Model on audience emotional responses.

3.11.6 Evaluation of the Audio Voice Model

1) Analysis of Expert Evaluations on the Suitability of the Audio Voice Model

Mean and standard deviation were used to determine the suitability ratings given by experts.

2) Analysis of Qualitative Feedback

Content analysis was conducted to summarize expert and audience suggestions for improving the Audio Voice Model.

3.12 Statistics used in research

3.12.1 Descriptive Statistics: Descriptive statistics summarized the basic characteristics of the collected data and presented it in a comprehensible manner.

1) Mean: The mean was calculated to describe the average score of the audience's positive emotions.

2) Standard Deviation: The standard deviation measured the variation and dispersion in the data.

3) Frequency: Frequency analysis determined the number of listeners categorized by gender or type of podcast host's voice.

4) Percentage: Percentages were used to show proportions, such as the gender distribution of participants.

3.12.2 Inferential Statistics: Inferential statistics analysed the data to test the research hypotheses.

1) Paired t-test:

1.1) The paired t-test compared the positive emotion scores of the audience before applying audio model and after applying audio model to the podcast.

1.2) It tested whether the stimuli (host’s voice and content) significantly influenced the audience's positive emotions.

2) ANOVA (Analysis of Variance):

2.1) ANOVA examined the differences in positive emotion scores between the group listening to male hosts and the group listening to female hosts.

2.2) It tested the effects of the host's gender and voice on the audience's positive emotions.

3) Pearson Correlation:

3.1) The Pearson correlation analysis determined the relationship between the audio voice model and the audience’s positive emotion scores.

3.2) It tested the correlation between the host's gender and the audience's positive emotion scores.

3.12.3 Microstate Analysis

Microstate analysis is used to detect and classify transient patterns of brain activity (attention allocation, emotional response intensity, emotional and cognitive integration, and changes in introspective activity) in the pre- and post-listening stages of podcast listening.

CHAPTER IV

Results

The title of the research is " Audio Voice Model to Reflect Audience Positive Emotion for Podcasting". The researcher has elaborated the findings of the study according to the research objectives as follows:

4.1 The impact of the voices and genders of podcast hosts on the audience's positive emotional responses.

4.2 Outcome of developing an audio voice model to reflect the audience's positive emotions for podcasting

4.3 The correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts.

4.4 The results of the compare the positive emotions of the audience before and after listening to a podcast.

Results

The research findings on the Audio Voice Model to Reflect Audience Positive Emotion for Podcasting are presented by the researcher as follows:

4.1 The impact of the voices and genders of podcast hosts on the audience's positive emotional responses

Table 4. 1 Basic Information of the Audience

Basic Information	Frequency	Percentage
1. Gender		
Male	25	50
Female	25	50
Other	0	0
Total	50	100

Basic Information	Frequency	Percentage
2. Age		
Under 18	1	2
18-25 years old	16	32
26-33 years old	15	30
34-41 years old	6	12
42-49 years old	8	16
50 years old and above	4	8
Total	50	100
3. Podcast Listening Frequency (per week)		
Every day	4	8
1 day	12	24
2 days	13	26
3 days	8	16
4 days	6	12
5 days	4	8
6 days	3	6
Total	50	100
4. Preferred Gender of Podcast Host		
Male	13	26
Female	32	64
No preference	5	10
Total	50	100

Form the Table 4.1 The audience's age distribution is skewed toward younger demographics, with 64% of listeners under 33 years old. Specifically, the largest age group is 18-25 years (32% of the total sample), followed closely by the 26-33 years group, which accounts for 30% of the audience. Meanwhile, older age groups are

represented to a lesser degree, with only 8% being 50 years or older and just 2% under 18. The youthful skew of this audience could influence emotional responses to podcast hosts, as younger audiences may respond differently to voices and gender representations compared to older groups. Younger listeners might be more likely to favor engaging, energetic voices or specific tones, while older listeners could prefer maturity and depth in a host's voice. These differences warrant further investigation.

The data shows a wide range in listening frequency, with only 8% of respondents listening to podcasts every day and most listening either 1 day (24%) or 2 days per week (26%). Less frequent listening patterns, with 3-5 days per week, account for 36% of the sample. Notably, 16% of the respondents listen to podcasts three times a week, suggesting that these listeners have adopted a regular but not daily routine. Understanding the correlation between frequency of podcast consumption and the impact of host voice and gender may reveal how often listeners' emotional responses are activated or sustained by auditory elements.

More frequent listeners may develop stronger preferences or emotional bonds with certain voices, while less frequent listeners might not exhibit as pronounced emotional responses. The majority of listeners, 64%, show a preference for female podcast hosts, while 26% prefer male hosts. Only 10% expressed no preference. This significant tilt toward female hosts may suggest that listeners find female voices more soothing, engaging, or emotionally resonant. Since the emotional responses of the audience are the focus of the study, the fact that a large portion of listeners prefers female hosts could have important implications for understanding how gender and vocal qualities contribute to emotional appeal. For instance, this preference might stem from cultural or psychological factors that associate female voices with warmth, empathy, or approachability. Alternatively, the preferences could be tied to the content of the podcasts they regularly consume, as certain genres may favor female hosts or female-led storytelling styles.

The provided statistics on audience demographics and preferences offer valuable insights for studying the relationship between podcast hosts’ voices, their genders, and the emotional responses of listeners. With a balanced gender representation, a youthful age profile, varied listening habits, and a strong preference for female hosts,

the dataset provides a solid foundation for examining how these factors influence audience engagement.

Table 4. 2 The results of the survey on key components of the audio voice model for reflecting positive emotions.

Components of a Audio Voice Model	Opinion level n=50		Meaning
	Mean	Std.	
1. Input Audio Processing	4.96	0.20	Highest
2. Audio Analysis	4.92	0.27	Highest
3. Emotion Adaptive Voice Synthesis	4.92	0.27	Highest
4. Creative Voice Modulation	4.90	0.30	Highest
5. Output Audio Generation	4.90	0.30	Highest
6. Evaluation and Optimization	4.94	0.24	Highest
7. Real-Time Listener Feedback Integration	4.92	0.27	Highest
8. Context-Aware Voice Adjustment	4.94	0.24	Highest
9. Multi-Language Emotional Mapping	4.22	0.79	Moderate
10. Emotional Continuity Enhancement	4.14	0.86	Moderate
11. Advanced Customization Dashboard	4.12	0.80	Moderate
12. Data-Driven Improvement Module	4.10	0.81	Moderate
Total	4.53	0.72	Highest

Research Results on the Key Components of the Audio Voice Model for Reflecting Positive Emotions The survey analysed expert opinions regarding the key components of an Audio Voice Model for reflecting positive emotions. The findings, summarized in Table 2, reveal insightful results about the essential elements required

for the model's design, with the overall evaluation indicating the highest suitability (Mean=4.53, Std.=0.72).

1. Key Findings:

1.1) Input Audio Processing: This component received the highest evaluation (Mean=4.96, Std.=0.20), indicating its critical importance in ensuring high-quality initial audio input for the model.

1.2) Audio Analysis: Experts rated this component highly (Mean=4.92, Std.=0.27), highlighting its role in accurately identifying emotional cues and nuances within the audio data.

1.3) Emotion Adaptive Voice Synthesis: This component was also ranked at the highest level (Mean=4.92, Std.=0.27). It emphasizes the importance of dynamically adapting voice characteristics to match the emotional intent of the content.

1.4) Creative Voice Modulation: This component achieved a high rating (Mean=4.90, Std.=0.30), underscoring its relevance in introducing creative tonal and pitch variations to enhance listener engagement.

1.5) Output Audio Generation: Rated equally high (Mean=4.90, Std.=0.30), this component reflects the necessity of generating high-quality audio outputs that resonate emotionally with listeners.

1.6) Evaluation and Optimization: This component also achieved a top rating (Mean=4.94, Std.=0.24), emphasizing the importance of incorporating systematic evaluation and feedback loops to ensure the model's continuous improvement.

1.7) Real-Time Listener Feedback Integration : This component was also ranked at the highest level (Mean=4.92, Std.=0.27). Enables podcast audio to dynamically adjust voice characteristics to optimize the emotional experience for the listener. This mechanism ensures that the audio consistently conveys positive emotions across different audience segments, dramatically increasing immersion and interactivity.

1.8) Context-Aware Voice Adjustment: This component also achieved a top rating (Mean=4.94, Std.=0.24), Enables podcast audio to adjust to content, listener context and real-time feedback, ensuring that the voice expression is highly matched

to the listener's emotional expectations. This feature relies on deep learning algorithms and real-time contextual analysis to accurately recognize contextual emotional needs.

2. Additional Insights :

While the primary components received the highest ratings, the survey also evaluated supplementary elements:

2.1) Components like Emotional Continuity Enhancement (Mean=4 . 1 4 , Std.=0 . 8 6) and Advanced Customization Dashboard (Mean=4 . 1 2 , Std.=0 . 8 0) demonstrated moderate importance for enhancing user adaptability.

2.2) The Data-Driven Improvement Module (Mean=4 . 1 0 , Std.=0 . 8 1) highlights the need for continuous learning to refine the model.

Based on the research findings, although the primary components received the highest scores, the survey also evaluated additional elements as opportunities for improvement to enhance the model's versatility. The researcher suggested that these supplementary components should be integrated with the six core evaluation criteria during the model development process to ensure the audio voice model is comprehensive and effective.

Table 4. 3 The Impact of the Host's Gender on the Audience's Positive Emotional Responses

No.	**Positive Emotional Lists**	**Positive Emotional Level** Female host voice **n=50**			**Positive Emotional Level** Male host voice **n=50**			**Difference**
		Mean	**Std.**	**Meaning**	**Mean**	**Std.**	**Meaning**	
1	The pitch of the host's voice makes	3.46	0.58	Very much	3.86	0.57	Very much	0.40

No.	Positive Emotional Lists	Positive Emotional Level — Female host voice n=50 Mean	Std.	Meaning	Positive Emotional Level — Male host voice n=50 Mean	Std.	Meaning	Difference
	me feel joyful							
2	The volume of the host's voice makes me feel excited	4.20	0.40	Very much	4.54	0.50	Very much	0.34
3	The rhythm of the host's speech enhances my sense of contentment	3.78	0.62	Very much	4.22	0.84	Very much	0.44
4	The timbre of the host's voice makes me feel loved and warm	4.15	0.43	Very much	3.58	0.50	Very much	0.57
5	The pauses in the host's speech make me feel hopeful	4.30	0.84	Very much	4.02	0.65	Very much	0.28
6	The speech rate of the host makes me feel grateful	4.52	0.50	Very much	4.06	0.24	Very much	0.46

No.	Positive Emotional Lists	Positive Emotional Level — Female host voice n=50 Mean	Std.	Meaning	Positive Emotional Level — Male host voice n=50 Mean	Std.	Meaning	Difference
7	The accent of the host reinforces my sense of passion	4.38	0.49	Very much	3.88	0.33	Very much	0.50
	Total	4.24	0.61	Very much	4.02	0.61	Very much	0.21

Results from Table 4.3: The Impact of the Host's Gender on the Audience's Positive Emotional Responses. The study evaluated the effect of the voices and genders of podcast hosts on the audience's positive emotional responses. The findings are summarized as follows:

1. Pitch of the Host's Voice

The pitch of the female host's voice made participants feel joyful (Mean = 3.46, SD = 0.58), while the male host's voice scored slightly higher (Mean = 3.86, SD = 0.57). The difference of 0.40 indicated a marginally greater impact of the male host's pitch on joy.

2. Volume of the Host's Voice

The volume of the female host's voice excited participants (Mean = 4.20, SD = 0.40), but the male host's voice elicited a slightly higher level of excitement (Mean = 4.54, SD = 0.50), with a difference of 0.34.

3. Rhythm of the Host's Speech

The rhythm of the female host's speech enhanced contentment (Mean = 3.78, SD = 0.62), but the male host's rhythm had a stronger impact (Mean = 4.22, SD = 0.84), with a difference of 0.44.

4. Timbre of the Host's Voice

The timbre of the female host's voice evoked a stronger sense of being loved and warm (Mean = 4.15, SD = 0.43) compared to the male host (Mean = 3.58, SD = 0.50), with a difference of 0.57 favoring the female host.

5. Pauses in the Host's Speech

The pauses used by the female host made participants feel more hopeful (Mean = 4.30, SD = 0.84) than the pauses in the male host's speech (Mean = 4.02, SD = 0.65), with a difference of 0.28.

6. Speech Rate of the Host

The female host's speech rate generated greater gratitude (Mean = 4.52, SD = 0.50) compared to the male host (Mean = 4.06, SD = 0.24), with a difference of 0.46.

7. Accent of the Host

The female host's accent reinforced a higher sense of passion (Mean = 4.38, SD = 0.49) than the male host (Mean = 3.88, SD = 0.33), with a difference of 0.50.

The overall positive emotional response was slightly higher for female hosts (Mean = 4.24, SD = 0.61) than for male hosts (Mean = 4.02, SD = 0.61). The total difference of 0.21 indicates a marginally stronger emotional impact from female hosts.

The findings reveal that both female and male hosts had significant impacts on positive emotional responses, with female hosts excelling in evoking feelings of being loved, grateful, and passionate, while male hosts had a stronger effect on joy, excitement, and contentment. This indicates that the emotional response varies by gender and specific vocal characteristics

Table 4. 4 Audio analysis unit for male and female podcast programs

The analysis of the physical sound characteristics and perceptual sound traits of podcast hosts of different genders (F_1-F_5 for females and M_1-M_5 for males) affected the 7 types of positive emotional responses from listeners (cheerfulness, excitement, contentment, warmth and love, hopefulness, gratitude, and passion). The research findings were as follows:

No.	Physical Sound Components	Units	Female					Male					Positive Emotion
			F1	F2	F3	F4	F5	M 1	M 2	M 3	M 4	M 5	
1	Amplitude	(Relative Intensity)	1.0	1.0	1.0	1.0	0.816	0.845	0.91	0.829	0.889	1.0	Joyful
2	Frequency	(Hz - Hertz)	187.63	264.55	169.45	315.89	335.79	168.67	225.02	330.42	299.15	45.98	Excited
3	Phase	(Radians)	-0.0	-3.14	-0.0	-3.14	2.856	-1.604	1.968	0.125	1.183	-3.14	Contentment
4	Speed	(BPM - Beats Per Minute)	121.79	146.13	113.07	152.43	198.18	107.5	101.67	165.38	193.88	117.43	Loved and Warm
5	Wavelength	(ms - Milliseconds)	255.82	166.69	260.25	139.63	270.32	912.33	603.12	603.61	701.89	958.92	Hopeful
6	Harmonic	(Spectral Centroid - Hz)	115.391	109.631	173.408	74.791	54.134	149.135	156.365	128.252	126.179	61.693	Grateful

7	Sharpness	(Sharpness Index)	0.87	1.57	0.219	0.28	0.549	0.261	1.18	0.138	1.339	1.129	Passion

From Table 4.4, the analysis results of the numerical values for the physical characteristics of sound and the perceived characteristics of the voices of podcast hosts of different genders are presented as follows.

1.Amplitude Analysis

Amplitude represents the relative strength of the sound, in this data, all female audio (F1-F4) has an amplitude of 1.0, while F5 is slightly lower at 0.816. Male audio (M1-M5) has an amplitude range of 0.829 - 1.0, with M3 being the lowest (0.829) and M5 being the highest (1.0). Overall, the amplitudes were more consistent for females and more variable for males. The data related to Joyful mood suggests that higher amplitude audio is more likely to trigger feelings of pleasure.

2.Frequency Analysis

The main frequency range of female audio was 169.45 Hz - 335.79 Hz, with F5 being the highest and F3 the lowest. Male audio ranges from 45.98 Hz - 330.42 Hz, with M5 being the lowest at 45.98 Hz. Higher frequency audio (such as F5 and M3) is usually associated with an Excited mood, whereas lower frequency audio (such as M5) may be associated with a more subdued or relaxed feeling.

3.Phase Analysis

The phase data ranges from -3.14 to 2.856, with most of the audio having a stable phase, such as F1 and F3 (both 0.0), but extreme phase (-3.14) occurs in F2, F4, and M5. Large phase variations (e.g., M2, M4) may affect the clarity and rhythmicity of the sound. The data suggests that more stable phases may be more consistent with Contentment, while more fluctuating phases may enhance the expressiveness of the sound.

4.Speed of Speech Analysis

The speed of speech ranged from 101.67 BPM to 198.18 BPM, with F5 and M4 having the highest speeds (198.18 BPM and 193.88 BPM) and M2 having the lowest (101.67 BPM). Faster speech rates (F5, M4) tend to bring excitement or energy, while slower speech rates (M2) may lead to a Loved and Warm or Calm

mood. The speed of speech affects the listener's perception of rhythm and determines the impact of the sound.

5.Wavelength Analysis

Wavelength (ms) is inversely proportional to frequency, ranging from 139.6 ms (F4) to 958.92 ms (M5). F4 has the shortest wavelength, which indicates that it is the sharpest and clearest sound, while M5 has the longest wavelength, which indicates that it is lower and softer. Shorter wavelengths are more likely to convey hopeful and energetic emotions, while longer wavelengths are usually associated with calm, steady emotions.

6.Harmonic Analysis

Harmonic frequencies range from 54.134 Hz (F5) to 173.408 Hz (F3), with F3 being the highest, and M5 and F5 the lowest (< 60 Hz). Higher harmonics usually enhance the brightness of the sound and make it more penetrating, while lower harmonics may make the sound sound softer. Data shows that sounds with higher harmonics are more likely to evoke a Grateful mood.

7.Sharpness Analysis

The Sharpness index ranges from 0.138 (M3) to 1.57 (F2), with F2 being the highest and M3 the lowest. Higher sharpness (e.g., F2, M5) makes the sound sharper and more prominent, possibly related to Passion, while lower sharpness (e.g., M3, F3) tends to be more moderate and stable. Higher sharpness increases the excitement of the sound and makes it more expressive.

Table 4. 5 Guidelines for Adjusting Sound Components by Podcast Host Gender to Enhance Positive Emotions

Positive Emotion	Amplitude (Relative Intensity)	Frequency (Hz)	Phase (Radians)	Speed (BPM)	Wavelength (ms)	Harmonic (Hz - Spectral Centroid)	Sharpness (Sharpness Index)
Joyful	Increase (0.9 - 1.0)	Moderate (170-265 Hz)	Stable (0.0 - 1.5)	Moderate (120-150 BPM)	Medium (166-260 ms)	Increase (110-170 Hz)	Moderate (0.5 - 1.0)
Excited	Increase (0.8 - 1.0)	High (250-335 Hz)	Varied (-3.14 to 2.85)	Fast (150-200 BPM)	Short (140-200 ms)	Increase (130-173 Hz)	High (1.2 - 1.5)

Contentment	Moderate (0.8 - 0.9)	Low (150-190 Hz)	Stable (0.0 - 1.5)	Slow (100-120 BPM)	Long (250-600 ms)	Decrease (100-140 Hz)	Low (0.2 - 0.5)
Loved & Warm	Moderate (0.8 - 0.9)	Balanced (190-250 Hz)	Stable (0.0 - 1.2)	Slow (100-120 BPM)	Long (600-900 ms)	Moderate (120-160 Hz)	Low (0.2 - 0.5)
Hopeful	Moderate (0.8 - 1.0)	High (270-330 Hz)	Varied (-1.5 to 2.0)	Fast (140-180 BPM)	Short (140-250 ms)	Moderate (120-160 Hz)	Moderate (0.5 - 1.0)
Grateful	Moderate (0.8 - 0.9)	Moderate (200-300 Hz)	Stable (0.0 - 1.5)	Moderate (100-130 BPM)	Medium (250-700 ms)	High (130-170 Hz)	Low (0.2 - 0.5)
Passion	High (0.9 - 1.0)	High (250-335 Hz)	Varied (-3.14 to 3.14)	Fast (150-200 BPM)	Short (140-260 ms)	High (150-175 Hz)	High (1.0 - 1.5)

Table 4.5 presents research findings that build upon the results in Table 4.4 to establish guidelines for adjusting physical sound components to enhance positive emotions in listeners. The findings highlight two key aspects:

1. Adjusting physical audio components to improve the overall sound quality of the podcast.

1.1 Amplitude (Relative Intensity)

1.1.1 High (0.9 - 1.0): Joyful, Excited, Passion (More energy and presence).

1.1.2 Moderate (0.8 - 0.9): Contentment, Loved & Warm, Grateful, Hopeful (Maintains warmth).

1.2 Frequency (Hz)

1.2.1 High (250-335 Hz): Excited, Passion, Hopeful (Brighter, engaging tones).

1.2.2 Moderate (200-300 Hz): Joyful, Grateful, Loved & Warm (Balanced warmth).

1.2.3 Low (150-190 Hz): Contentment (Deep, soothing tone).

1.3 Phase (Radians)

1.3.1 Stable (0.0 - 1.5): Joyful, Contentment, Loved & Warm, Grateful (Smooth, natural flow).

1.3.2 Varied (-3.14 to 3.14): Excited, Passion, Hopeful (Expressive, lively transitions).

1.4 Speed (BPM)

1.4.1 Fast (150-200 BPM): Excited, Passion (High-energy, dynamic).

1.4.2 Moderate (120-150 BPM): Joyful, Hopeful (Lively, but controlled).

1.4.3 Slow (100-120 BPM): Contentment, Loved & Warm, Grateful (Relaxed, soothing).

1.5 Wavelength (ms)

1.5.1 Short (140-250 ms): Excited, Hopeful, Passion (Sharp, clear tones).

1.5.2 Medium (250-600 ms): Joyful, Grateful, Contentment (Balanced clarity).

1.5.3 Long (600-900 ms): Loved & Warm (Soft, warm resonance).

1.6 Harmonic (Spectral Centroid - Hz)

1.6.1 High (150-175 Hz): Passion, Excited (More energy, resonance).

1.6.2 Moderate (120-160 Hz): Hopeful, Grateful, Joyful (Balanced harmonic depth).

1.6.3 Low (100-140 Hz): Contentment, Loved & Warm (Softer, relaxed).

1.7 Sharpness (Sharpness Index)

1.7.1 High (1.0 - 1.5): Passion, Excited (More expression and presence).

1.7.2 Moderate (0.5 - 1.0): Joyful, Hopeful (Clear but controlled).

Low (0.2 - 0.5): Contentment, Loved & Warm, Grateful (Smooth, calm sound).

2. Modifying physical audio components to optimize effectiveness based on the gender of the podcast host, distinguishing between male and female voices.

Table 4. 6 Female podcast host sound optimization for audience engagement

Positive Emotion	Amplitude	Frequency (Hz)	Phase (Radians)	Speed (BPM)	Wavelength (ms)	Harmonic (Hz)	Sharpness (Index)
Joyful	0.9 - 1.0	170 - 265	0.0 - 1.5	120 - 150	166 - 260	110 - 170	0.5 - 1.0
Excited	0.8 - 1.0	250 - 335	-3.14 to 2.85	150 - 200	140 - 200	130 - 173	1.2 - 1.5
Contentment	0.8 - 0.9	150 - 190	0.0 - 1.5	100 - 120	250 - 600	100 - 140	0.2 - 0.5
Loved & Warm	0.8 - 0.9	190 - 250	0.0 - 1.2	100 - 120	600 - 900	120 - 160	0.2 - 0.5
Hopeful	0.8 - 1.0	270 - 330	-1.5 to 2.0	140 - 180	140 - 250	120 - 160	0.5 - 1.0
Grateful	0.8 - 0.9	200 - 300	0.0 - 1.5	100 - 130	250 - 700	130 - 170	0.2 - 0.5
Passion	0.9 - 1.0	250 - 335	-3.14 to 3.14	150 - 200	140 - 260	150 - 175	1.0 - 1.5

From Table 4.6, the research findings indicated that adjusting the physical sound components of female podcast hosts effectively enhanced positive emotional responses from the audience. By increasing or decreasing amplitude, frequency, phase, speed, wavelength, harmonic, and sharpness, different emotional effects were achieved.

1. Joyful and excited emotions required increased amplitude, frequency, and speech speed, along with a shortened wavelength to create an energetic and engaging vocal tone.

2. Contentment and warmth were best conveyed through stable phase, slower speech speed, and reduced sharpness, ensuring a soothing and friendly listening experience.

3. Hopeful and passionate emotions were enhanced by raising frequency and speech speed, while dynamic phase adjustments contributed to an expressive and inspiring vocal delivery.

In conclusion, optimizing sound components effectively improved vocal communication, enabling female podcast hosts to enhance positive emotions and engage their audience more effectively.

Table 4. 7 Male podcast host sound optimization for audience engagement

Positive Emotion	Amplitude	Frequency (Hz)	Phase (Radians)	Speed (BPM)	Wavelength (ms)	Harmonic (Hz)	Sharpness (Index)
Joyful	0.8 - 1.0	150 - 265	0.0 - 1.5	120 - 140	250 - 600	110 - 170	0.5 - 1.0
Excited	0.8 - 1.0	250 - 330	-3.14 to 2.85	150 - 190	140 - 220	125 - 170	1.0 - 1.5
Contentment	0.8 - 0.9	150 - 190	0.0 - 1.5	100 - 120	300 - 650	100 - 140	0.2 - 0.5
Loved & Warm	0.8 - 0.9	180 - 250	0.0 - 1.2	100 - 120	650 - 950	110 - 150	0.2 - 0.5
Hopeful	0.8 - 1.0	250 - 315	-1.5 to 2.0	130 - 170	200 - 350	115 - 155	0.5 - 1.0
Grateful	0.8 - 0.9	180 - 290	0.0 - 1.5	100 - 130	300 - 750	120 - 165	0.2 - 0.5
Passion	0.9 - 1.0	250 - 335	-3.14 to 3.14	150 - 200	150 - 275	145 - 175	1.2 - 1.5

Table 4.7 the research findings indicated that adjusting the sound components of male podcast hosts effectively enhanced audience engagement and promoted positive emotions.

1. Joyful and excited emotions were best conveyed by increasing amplitude, frequency, and speech speed, while reducing wavelength created a more energetic and engaging vocal tone.

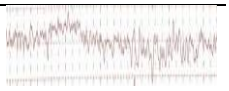
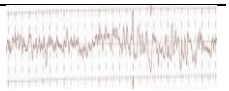
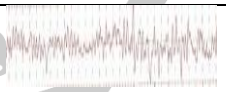
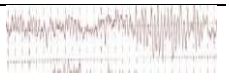
2. Contentment and warmth required stable phase, slower speech speed, and reduced sharpness, ensuring a steady and approachable listening experience.

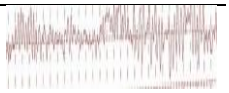
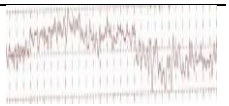
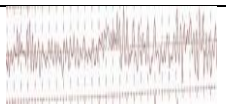
3. Hopeful and passionate emotions were strengthened by raising frequency and speech speed, while dynamic phase adjustments contributed to a deeper and more intense vocal delivery.

In conclusion, optimizing sound components effectively improved vocal communication, allowing male podcast hosts to enhance emotional expression and maximize audience engagement.

Highlight how different podcast audio frequencies influence emotional states through distinct EEG patterns.

Table 4. 8 The Relationship Between Podcast Audio Frequency and EEG Emotional Responses

Emotion Category	Audio Frequency (Hz)	EEG Band	Brainwave State	Audio Characteristics	Waveform Image
Joyful	187.63 Hz	Beta (13-30 Hz)	Happiness, Vitality	High-frequency vibration, concentrated energy	
Excited	264.55 Hz	High Alpha / Low Beta	Excitement, Focus	Large amplitude variations, enhanced oscillation	
Contentment	169.45 Hz	Alpha (8-13 Hz)	Relaxation, Satisfaction	Long wavelengths, even amplitude	
Love	315.89 Hz	Theta (4-8 Hz)	Intimacy, Calmness	Soft curves, low-frequency modulation	

Hopeful	335.79 Hz	Low Beta / Alpha	Positivity, Expectation	Gradually increasing oscillations, mid-frequency features	
Grateful	173.41 Hz	Alpha (8-12 Hz)	Meditation, Tranquility	Slight fluctuations, gentle variations	
Passion	225.02 Hz	Gamma (30+ Hz)	High Excitement	Strong amplitude oscillations, high-frequency characteristics	

The results from Table 4.8 show the relationship between podcast audio frequency and EEG emotional response as follows:

1. Joyful Audio and Beta Wave Activation

The corresponding audio frequency of Joyful is 187.63 Hz, and its EEG is mainly active in the Beta frequency band (13-30 Hz), which is related to the emotional state of being happy and energetic. The audio waveform shows that the amplitude changes are more uniform, with moderate fluctuations, showing smooth and rhythmic characteristics. This type of audio usually has a certain high-frequency component, which makes the sound clearer and brighter, and can cause a pleasant psychological experience for the listener.EEG data show that when listening to this type of audio, the energy level of the Beta band is increased, which indicates that the brain is in an active and excited state, and the activity in the Alpha band is reduced, which indicates that the listener's attention is focused on the sound stimulus, and the perception of pleasure is strengthened.

2. Excited Audio and Alpha-Beta Stimulation

The audio frequency corresponding to Excited is 264.55 Hz, and the EEG mainly shows the enhancement of high Alpha (8-13 Hz) and low Beta (13-20 Hz), which indicates that the listener is in the state of concentration, excitement, and positive mental state. The audio waveform graphs show large changes in amplitude, significantly enhanced fluctuation

amplitude, and higher waveform density, implying that the audio signals are more energetic and rhythmic. This characteristic makes the audio more penetrating, which can effectively stimulate the auditory system and trigger the listener's excitement. EEG data further shows that the high end of the Alpha band and the low end of the Beta band are significantly enhanced when listening to excited audio, implying that the brain is balancing excitement and maintaining attention, and also demonstrating a certain psychological arousal effect.

3. Contentment Audio and Alpha Wave Enhancement

Contentment corresponds to an audio frequency of 169.45 Hz and mainly affects the Alpha band (8-13 Hz), which is usually associated with relaxation, satisfaction and comfort. The waveform shows that the audio signal is smooth, with small amplitude variations and a gentle curve, indicating lower energy and a slower tempo. This audio characteristic makes the sound more soothing and less irritating, making it easier for the listener to enter a calm emotional state. EEG studies have shown that when listening to this audio, an increase in the Alpha band indicates a relaxation of the brain and a lowering of the stress level, and an increase in the Theta band (4-8 Hz), which is often associated with contemplation, meditation, and deep relaxation.

4. Loved Audio and Theta Wave Activation

Loved and Warm corresponds to an audio frequency of 315.89 Hz, and the EEG is dominated by an increase in the Theta band (4-8 Hz), which is associated with feelings of intimacy, emotional connection, and relaxation. The audio waveform graph shows a more uniform amplitude with smooth changes and a gentle rhythm, and the waveform shows a slight progressive enhancement, giving the sound a more natural smoothness. This type of audio modulation creates a warmer, more comfortable listening environment, and gives the listener a sense of intimacy and security. eeg found that when listening to this audio, the enhancement of the Theta band reflects the brain's ability to enter into a more relaxed and immersive situation, and also activates parts of the Alpha band, which supports a warm, reassuring psychological experience.

5. Hopeful Audio and Alpha-Beta Balance

Hopeful corresponds to an audio frequency of 335.79 Hz, and the EEG is mainly distributed in the low Beta (13-20 Hz) and Alpha (8-13 Hz), indicating that the listener is in an emotional state of positive anticipation. The waveform graph shows that the amplitude variation is slightly enhanced but still remains smooth, implying that the sound has a

moderate sense of fluctuation so that it is neither too rapid nor too flat. This audio feature is suitable for enhancing the atmosphere of confidence and hope, allowing listeners to feel motivated and optimistic. EEG studies have shown a slight increase in activity in the Alpha band and activation in the low Beta band when listening to Hopeful audio, suggesting that the brain maintains a sense of relaxation while maintaining a degree of cognitive activity, allowing the listener to remain attentive and stable in a hopeful mood.

6. Grateful Audio and Deep Alpha Relaxation

Grateful corresponds to an audio frequency of 173.41 Hz, and the EEG mainly shows enhancement in the Alpha band (8-12 Hz), which is associated with meditation, introspection, and peace of mind. The waveform graph shows that the audio signal is less fluctuating, with gentle changes and an even, soft sound, making it easier to bring about a relaxing and calming mental experience. This audio signature is often used in meditative, prayerful, or deep-thinking situations, and can help the listener to regulate their emotions and become more at peace with themselves. eeg has found that when listening to Grateful audio, there is a significant enhancement in the Alpha band, and there may be a certain amount of enhancement in the Theta band, which is often associated with relaxation, increased gratitude, and the establishment of inner peace.

7. Passion Audio and Gamma Wave Boost

Passion corresponds to an audio frequency of 225.02 Hz, and the EEG mainly shows activation of the Gamma band (30+ Hz), which is usually associated with high excitement, complex thinking and strong emotions. The waveform graph shows that the amplitude of the audio signal changes drastically, and the waveform is dense and highly concentrated in energy, showing a strong oscillation pattern, making the sound more explosive and expressive. This audio characteristic allows the listener to quickly enter an emotional state of excitement and motivation. EEG studies have found that when listening to Passion audio, the significant enhancement of the Gamma band reflects the high level of brain activity and emotional mobilization, and that the high end of the Beta band may also be affected, further enhancing the perception of passion and motivation and allowing the listener to enter a state of heightened concentration and emotional intensity more easily. highly focused and emotionally charged state.

4.2 Outcome of developing an audio voice model to reflect the audience's positive emotions for podcasting

This section presents a qualitative analysis based on expert evaluations of an audio voice model designed to convey positive emotions in podcast content. 5 experts from various fields voice coaching, podcast production, communication, audio engineering, and academia provided feedback on the model's effectiveness and suggested improvements. Their perspectives offer a comprehensive understanding of the model's strengths in emotional engagement and areas for enhancement, which could improve listener resonance and emotional depth.

4.2.1 The audio voice model for reflecting audience positive emotions for podcasting.

AVM-PE (Audio Voice Model for Positive Emotions), developed by the researcher, was derived through the study, analysis, and synthesis of principles, theories, relevant research, and expert interviews. The research findings identified 6 components as follows.

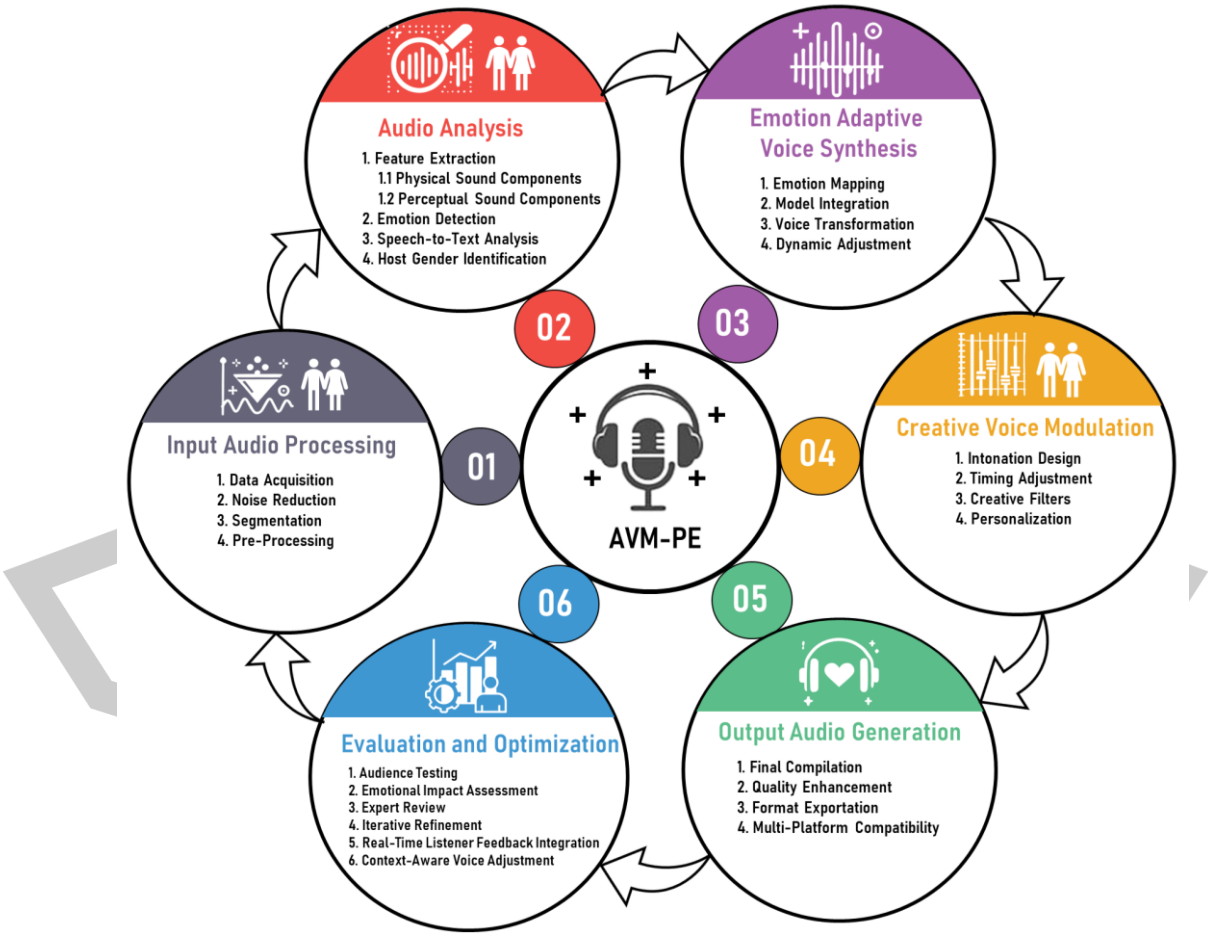


Figure 4. 1 AVM-PE (Audio Voice Model for Positive Emotions)

From Figure 4 .1, the detailed description of the 6 components of the AVM-PE (Audio Voice Model for Positive Emotions) can be explained as follows to clarify the approach for applying the model in voice design.

1) Input Audio Processing

1.1) Data Acquisition: Collect audio data from podcast hosts in various formats.

1.2) Noise reduction: Ensures clear audio input by applying gain processing to the human voice and spillover processing (which can retain weak background sounds) to programs with background music.

1.3) Segmentation: Divide audio files into smaller, analyzable units based on time or frequency.

1 .4) Pre-Processing: Normalize the audio volume and prepare it for further analysis.

2) Audio Analysis

2.1) Feature Extraction: Identified key emotional signals in speech based on physical and perceptual sound components. The analysis used Audio Signal Processing and Analysis Tools to extract relevant features.

2.1.1) Physical Sound Components: Included amplitude, frequency, phase, speed, wavelength, harmonics, and sharpness.

2.1.2) Perceptual Sound Components: Comprised pitch, volume, speech patterns, rhythm, timbre, accent, and pauses.

2.2) Emotion Detection: This process involves analyzing audio characteristics to identify emotional cues, particularly 7 positive emotions, Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.

2.3) Speech-to-Text Analysis: Convert spoken words into text for additional emotional context.

2.4) Host Gender Identification: Match audio with host profiles to account for individual vocal characteristics.

3) Emotion Adaptive Voice Synthesis

3.1) Emotion Mapping: Map detected emotions to predefined positive emotional responses.

3.2) Model Integration: Integrate synthesized emotional responses into the audio.

3.3) Voice Transformation: Adapt vocal characteristics to enhance emotional resonance.

3.4) Dynamic Adjustment: Fine-tune the synthesis in real-time based on feedback loops.

4) Creative Voice Modulation

4.1) Intonation Design: Adjust pitch and tone to reflect the intended emotion.

4.2) Timing Adjustment: Modify pacing and pauses to enhance emotional impact.

4.3) Creative Filters: Apply vocal effects to increase listener engagement.

4.4) Personalization: Customize audio modulation settings to align with podcast host themes and gender.

5) Output Audio Generation

5.1) Final Compilation: Combine processed and modulated audio components into a single track.

5.2) Quality Enhancement: Apply final audio mastering techniques for professional quality.

5.3) Format Exportation: Convert the audio into podcast-ready formats (e.g., MP3, WAV).

5.4) Multi-Platform Compatibility: Ensure the audio is optimized for various streaming platforms.

6) Evaluation and Optimization

6.1) Audience Testing: Conduct trials with target listener groups to gather initial feedback on audio effectiveness.

6.2) Emotional Impact Assessment: Analyze how well the audio conveys intended emotions based on listener responses.

6.3) Expert Review: Seek evaluations from specialists in audio engineering and communication to ensure quality and effectiveness.

6.4) Iterative Refinement: Utilize feedback data to make systematic improvements to the voice model for enhanced performance.

6.5) Real-Time Listener Feedback Integration: Employ AI-driven analytics to assess audience engagement and dynamically adjust voice characteristics based on real-time reactions.

6.6) Context-Aware Voice Adjustment: Modify vocal tone, pitch, and tempo according to the content theme, audience demographics, and listening environment to optimize contextual relevance.

4.2.2 The suitability of the audio voice model to reflect the audience's positive emotions for podcasting

The results of the analysis of opinions from 5 experts regarding the suitability of the AVM-PE (Audio Voice Model for Positive Emotions) are presented by the researcher as follows:

Table 4. 9 Basic Information of the Expert

Basic Information	Frequency	Percentage
1. Gender		
Male	3	60
Female	2	40
Total	5	100
2. Age		
20-30 years old	1	20
31-40 years old	1	20
41-50 years old	3	60
51 years old and above	0	0
Total	5	100
3. Highest Educational Qualification		
Bachelor's Degree	1	20
Master's Degree	3	60
Doctoral Degree	1	20
Other (please specify)	0	0
Total	5	100
4. Field of Study		
Audio Engineering	2	40
Media Production	1	20
Communication Studies	1	20
Psychology	1	20

Basic Information	Frequency	Percentage
Other (please specify)	0	0
Total	5	100
5. Current Occupation		
Academic (Professor, Lecturer)	1	20
Audio Engineer	1	20
Media Producer	2	40
Podcaster	1	20
Other (please specify)	0	0
Total	5	100
6. Years of Experience in Audio Design, Podcasting, or Related Fields		
under 3 years	1	20
3-6 years	1	20
7-10 years	2	40
More than 10 years	1	20
Total	5	100

From Table 4.9, the panel of experts represents a balanced mix of genders, with the majority being male (60%).

Most experts are aged between 41-50 years (60%) and hold a Master's degree (60%).

Their professional backgrounds span diverse fields, including audio engineering, media production, communication studies, and psychology.

Two experts have 7-10 years of experience, indicating a strong level of expertise in audio design and podcasting-related fields.

This diverse and experienced group provides credible and valuable insights into evaluating the suitability of the Audio Voice Model for Positive Emotions.

Table 4. 10 Evaluation of the Suitability of the Audio Voice Model Components

No.	Audio Voice Model Components	Suitability level n=5		Meaning
		Mean	Std.	
1	**Input Audio Processing**			
	1.1. Data Acquisition: Collect audio data from podcast hosts in various formats.	4.80	0.41	Highest
	1.2. Noise Reduction: Filter out background noise to ensure clarity in audio inputs.	4.80	0.41	Highest
	1.3. Segmentation: Divide audio files into smaller, analyzable units based on time or frequency.	4.80	0.42	Highest
	1.4. Pre-Processing: Normalize the audio volume and prepare it for further analysis.	4.80	0.52	Highest
	Total	4.80	0.41	**Highest**
2	**Audio Analysis**			
	2.1) Feature Extraction: Identified key emotional signals in speech based on physical and perceptual sound components			
	2.2.1) Physical Sound Components: Included amplitude, frequency, phase, speed, wavelength, harmonics, and sharpness.	4.80	0.41	Highest
	2.2.2) Perceptual Sound Components: Comprised pitch, volume, speech patterns, rhythm, timbre, accent, and pauses.	4.80	0.41	Highest
	2.2. Emotion Detection: This process involves analyzing audio characteristics to identify emotional cues, particularly 7 positive emotions, Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.	4.80	0.41	Highest
	2.3. Speech-to-Text Analysis: Convert spoken words into text for additional emotional context.	4.80	0.42	Highest
	2.4. Host Gender Identification: Match audio with host profiles to account for individual vocal characteristics.	4.80	0.42	Highest

No.	Audio Voice Model Components	Suitability level n=5		Meaning
		Mean	Std.	
	Total	4.80	0.41	4.80
3	**Emotion Adaptive Voice Synthesis**			
	3.1. Emotion Mapping: Map detected emotions to predefined positive emotional responses.	4.80	0.41	Highest
	3.2. Model Integration: Integrate synthesized emotional responses into the audio.	4.80	0.41	Highest
	3.3. Voice Transformation: Adapt vocal characteristics to enhance emotional resonance.	4.80	0.42	Highest
	3.4. Dynamic Adjustment: Fine-tune the synthesis in real-time based on feedback loops.	4.80	0.42	Highest
	Total	4.80	0.41	**Highest**
4	**Creative Voice Modulation**			
	4.1. Intonation Design: Adjust pitch and tone to reflect the intended emotion.	4.80	0.44	Highest
	4.2. Timing Adjustment: Modify pacing and pauses to enhance emotional impact.	4.80	0.46	Highest
	4.3. Creative Filters: Apply vocal effects to increase listener engagement.	4.60	0.48	Highest
	4.4. Personalization: Customize audio modulation settings to align with podcast host themes and gender.	4.80	0.42	Highest
	Total	4.75	0.44	**Highest**
5	**Output Audio Generation**			
	5.1. Final Compilation: Combine processed and modulated audio components into a single track.	4.80	0.41	Highest
	5.2. Quality Enhancement: Apply final audio mastering techniques for professional quality.	4.80	0.41	Highest

No.	Audio Voice Model Components	Suitability level n=5		Meaning
		Mean	Std.	
	5.3. Format Exportation: Convert the audio into podcast-ready formats (e.g., MP3, WAV).	4.80	0.42	Highest
	5.4. Multi-Platform Compatibility: Ensure the audio is optimized for various streaming platforms.	4.80	0.42	Highest
	Total	4.80	0.41	**Highest**
6	**Evaluation and Optimization**			
	6.1. Audience Testing: Present the audio to a test group for initial feedback.	4.80	0.41	Highest
	6.2. Emotional Impact Assessment: Analyze how well the audio conveys positive emotions.	4.80	0.41	Highest
	6.3. Expert Review: Gather evaluations from experts in audio engineering and communication.	4.80	0.42	Highest
	6.4. Iterative Refinement: Use feedback to adjust and improve the model for future applications.	4.80	0.45	Highest
	6.5) Real-Time Listener Feedback Integration: Employ AI-driven analytics to assess audience engagement and dynamically adjust voice characteristics based on real-time reactions.	4.80	0.41	Highest
	6.6) Context-Aware Voice Adjustment: Modify vocal tone, pitch, and tempo according to the content theme, audience demographics, and listening environment to optimize contextual relevance.	4.80	0.42	Highest
	Total	4.80	0.41	**Highest**
	Overall Total	4.80	0.42	**Highest**

Table 4.10 summarizes expert evaluations of the 6 core components of the Audio Voice Model for Reflecting Audience Positive Emotions in Podcasting. The evaluations were based on a 5-point scale, with the overall indicating the model was

rated as "highest suitability" (Mean=4.80, Std.=0.42) across all components. Below are detailed results and highlights for each component.

1. Input Audio Processing

Noise reduction and segmentation processes were rated highest, emphasizing the model's ability to ensure clear and analyzable audio inputs the total result highest (Mean = 4.80, SD = 0.41)

2. Audio Analysis

Emotion detection and host identification achieved the highest scores, showcasing the model’s accuracy in identifying emotional and individual vocal characteristics. Feature extraction, although rated slightly lower (Mean = 4.40), received a recommendation to expand the criteria for analysis and was rated as most suitable (Mean = 4.70, SD = 0.47)

3. Emotion Adaptive Voice Synthesis

The model’s ability to dynamically adjust voice synthesis in real-time received universal praise for its capability to maintain emotional resonance and was rated as the most suitable (Mean = 4.80, SD = 0.41)

4. Creative Voice Modulation

Creative filters scored slightly lower (Mean = 4.60), with experts recommending incorporating A tools to enhance vocal effects and adjustments. Tone design and tempo adjustment were crucial to capturing the emotional differences presented by male and female podcast hosts and were best overall (mean = 4.75, standard deviation = 0.44).

5. Output Audio Generation

Experts unanimously rated this component as highly suitable, emphasizing the model’s versatility and readiness for professional multi-platform use, which overall were most suitable (Mean = 4.80, SD = 0.41)

6. Evaluation and Optimization

The iterative refinement process was praised for enabling continuous improvements to align with audience feedback and emotional impact goals and was rated as the most suitable (Mean = 4.80, SD = 0.41)

4.3 The correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts

4.3.1 Experimental Results of Using AVM-PE (Audio Voice Model for Positive Emotions) to Reflect Positive Emotions in Podcast Broadcasting

The experiment involved a sample group consisting of 5 male hosts, 5 female hosts, and 50 listeners. The study examined the effectiveness of the AVM-PE (Audio Voice Model for Positive Emotions) in enhancing positive emotional resonance in podcast broadcasting. In this experiment, 5 male and 5 female podcast hosts were trained to use the AVM-PE model to generate positive voice patterns for 7 different stories. Each story was designed to reflect a specific positive emotion: happiness, excitement, satisfaction, love and warmth, hope, gratitude, and passion. The findings revealed that:

Table 4. 11 Results of Testing the Positive Emotion Reflective Voice Model Application for Podcast Production

				7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners.																Difference
				1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		
No	Testing	Gender of Hosts	Audiences	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
1	Audio before applying model	Female	50	3.36	0.48	3.39	0.76	3.84	0.69	3.78	0.71	3.76	0.72	3.65	0.75	3.69	0.74	**3.64**	**0.69**	0.12
		Male	50	3.48	0.68	3.67	0.75	3.33	0.75	3.65	0.75	3.43	0.76	3.57	0.76	3.47	0.77	**3.51**	**0.75**	
		Total		**3.42**	**0.58**	**3.53**	**0.75**	**3.58**	**0.72**	**3.71**	**0.73**	**3.59**	**0.74**	**3.61**	**0.76**	**3.58**	**0.75**	3.58	**0.72**	
	Audio after applying model	Female	50	3.46	0.50	4.37	0.49	4.82	0.39	4.61	0.49	4.59	0.50	4.57	0.50	4.57	0.50	**4.43**	**0.48**	0.11
		Male	50	3.56	0.67	4.31	0.51	4.49	0.51	4.53	0.50	4.43	0.50	4.47	0.50	4.45	0.50	**4.32**	**0.53**	
		Total		3.51	0.59	4.34	0.50	4.65	0.45	4.57	0.50	4.51	0.50	4.52	0.50	4.51	0.50	**4.37**	**0.51**	
2	Audio	Female	50	3.54	0.50	3.73	0.73	3.73	0.73	4.29	0.54	4.45	0.54	3.65	0.75	3.51	0.77	**3.84**	**0.65**	0.09

No	Testing	Gender of Hosts	Audiences	**7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners.** 1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		Difference
				Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
	before applying model	Male	50	3.46	0.50	3.53	0.77	3.67	0.75	4.14	0.41	4.39	0.53	3.49	0.77	3.61	0.76	**3.76**	**0.64**	
		Total		3.50	0.50	3.63	0.75	3.70	0.74	4.21	0.47	4.42	0.54	3.57	0.76	3.56	0.76	**3.80**	**0.65**	
	Audio after applying model	Female	50	4.04	0.90	4.12	0.39	4.63	0.49	4.76	0.43	4.65	0.48	4.67	0.47	4.45	0.50	**4.48**	**0.52**	0.14
		Male	50	3.58	0.70	4.14	0.35	4.47	0.50	4.55	0.50	4.53	0.50	4.49	0.51	4.59	0.50	**4.34**	**0.51**	
		Total		3.81	0.80	4.13	0.37	4.55	0.50	4.65	0.47	4.59	0.49	4.58	0.49	4.52	0.50	**4.41**	**0.52**	
3	Audio before	Female	50	3.68	0.47	3.80	0.71	3.88	0.67	4.24	0.48	3.84	0.69	3.63	0.76	3.65	0.75	**3.82**	**0.65**	-0.09

No	Testing	Gender of Hosts	Audiences	7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners. 1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		Difference
				Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
	applying model	Male	50	3.78	0.42	3.63	0.76	4.18	0.60	4.65	0.52	4.29	0.54	3.45	0.77	3.39	0.76	**3.91**	**0.62**	
		Total		3.73	0.44	3.71	0.73	4.03	0.63	4.45	0.50	4.06	0.61	3.54	0.76	3.52	0.76	**3.86**	**0.63**	
	Audio after applying model	Female	50	4.82	0.39	4.22	0.42	4.37	0.49	4.69	0.47	4.20	0.41	4.65	0.48	4.73	0.45	**4.53**	**0.44**	0.06
		Male	50	4.68	0.47	4.06	0.24	4.55	0.50	4.53	0.50	4.53	0.50	4.45	0.50	4.45	0.50	**4.46**	**0.46**	
		Total		4.75	0.43	4.14	0.33	4.46	0.49	4.61	0.48	4.37	0.46	4.55	0.49	4.59	0.47	**4.50**	**0.45**	
4	Audio before applying model	Female	50	4.40	0.70	3.71	0.74	3.94	0.63	4.41	0.54	4.43	0.54	3.43	0.76	3.63	0.76	**3.99**	**0.67**	0.03
		Male	50	4.40	0.70	4.24	0.48	3.73	0.73	4.18	0.44	4.57	0.54	3.59	0.76	3.47	0.77	**4.03**	**0.63**	

No	Testing	Gender of Hosts	Audiences	7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners.																Difference
				1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		
				Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
		Total		4.40	0.70	3.98	0.61	3.84	0.68	4.30	0.49	4.50	0.54	3.51	0.76	3.55	0.76	**4.01**	**0.65**	
	Audio after applying model	Female	50	4.54	0.61	4.27	0.45	4.51	0.51	4.80	0.41	4.51	0.51	4.49	0.51	4.53	0.50	**4.52**	**0.50**	0.06
		Male	50	4.32	0.79	4.39	0.49	4.29	0.46	4.63	0.49	4.65	0.48	4.61	0.49	4.35	0.48	**4.46**	**0.53**	
		Total		4.43	0.70	4.33	0.47	4.40	0.48	4.71	0.45	4.58	0.49	4.55	0.50	4.44	0.49	**4.49**	**0.51**	
5	Audio before applying model	Female	50	4.58	0.78	4.18	0.67	4.37	0.53	4.20	0.46	4.41	0.54	4.29	0.50	3.76	0.72	**4.25**	**0.60**	0.09
		Male	50	4.62	0.78	3.86	0.74	4.67	0.52	4.24	0.48	4.18	0.44	4.08	0.49	3.47	0.77	**4.16**	**0.60**	
		Total		4.60	0.78	4.02	0.70	4.52	0.52	4.22	0.47	4.30	0.49	4.18	0.50	3.61	0.74	**4.21**	**0.60**	

No	Testing	Gender of Hosts	Audiences	1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		Difference
				Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
	7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners.																			
	Audio after applying model	Female	50	4.62	0.49	4.41	0.50	4.43	0.50	4.63	0.49	4.53	0.50	4.61	0.49	4.71	0.46	**4.56**	**0.49**	0.04
		Male	50	4.72	0.45	4.24	0.43	4.78	0.42	4.61	0.49	4.33	0.47	4.45	0.50	4.53	0.50	**4.52**	**0.47**	
		Total		4.67	0.47	4.33	0.47	4.60	0.46	4.62	0.49	4.43	0.49	4.53	0.50	4.62	0.48	**4.54**	**0.48**	
	Audio before applying model Female Total			**3.91**	**0.59**	**3.76**	**0.72**	**3.95**	**0.65**	**4.18**	**0.55**	**4.18**	**0.61**	**3.73**	**0.70**	**3.65**	**0.75**	**3.91**	**0.65**	**0.59**
	Audio after applying model Female Total			**4.30**	**0.58**	**4.28**	**0.45**	**4.55**	**0.47**	**4.70**	**0.46**	**4.50**	**0.48**	**4.60**	**0.49**	**4.60**	**0.48**	**4.50**	**0.49**	
	Audio before applying model Male Total			**3.95**	**0.62**	**3.79**	**0.70**	**3.92**	**0.67**	**4.18**	**0.52**	**4.17**	**0.56**	**3.64**	**0.71**	**3.48**	**0.76**	**3.87**	**0.65**	**0.55**

No	Testing	Gender of Hosts	Audiences	7 positive contents narrated by male and female hosts were presented to an audience of 50 listeners.																Difference
				1 Joyful		2 Excited		3. Contentment		4. Loved and warm		5. Hopeful		6. Grateful		7. Passion		**Overall Total**		
				Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	**Mean**	**Std.**	
	Audio after applying model Male Total			**4.17**	**0.62**	**4.23**	**0.41**	**4.51**	**0.48**	**4.57**	**0.50**	**4.49**	**0.49**	**4.49**	**0.50**	**4.47**	**0.50**	**4.42**	**0.50**	
	Audio before applying model Overall Total			**3.93**	**0.60**	**3.78**	**0.71**	**3.93**	**0.66**	**4.18**	**0.53**	**4.17**	**0.58**	**3.68**	**0.71**	**3.57**	**0.76**	**3.89**	**0.65**	**0..57**
	Audio after applying model Overall Total			**4.23**	**0.60**	**4.25**	**0.43**	**4.53**	**0.48**	**4.63**	**0.48**	**4.50**	**0.49**	**4.55**	**0.50**	**4.54**	**0.49**	**4.46**	**0.49**	

From Table 4.11, Results of Testing the Positive Emotion Reflective Voice Model for Podcast Production. The study aimed to evaluate the effectiveness of the Positive Emotion Reflective Audio Voice Model in enhancing positive emotions among listeners of a podcast. The research involved 50 participants who listened to 7 positive content episodes narrated by both male and female hosts. The listeners' emotions were assessed audio before applying model and audio after applying model to the podcasts across 7 categories: Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion. The results found that:

1. Joyful

1.1 Audio before applying model: The mean score for female listeners was 3.91 (Std. = 0.65) and for male listeners was 3.87 (Std. = 0.65). The total mean score was 3.89 (Std. = 0.65).

1.2 Audio after applying model: The mean score for female listeners increased to 4.50 (Std. = 0.49) and for male listeners to 4.42 (Std. = 0.50). The total mean score rose to 4.46 (Std. = 0.49).

2. Excited

2.1 Audio before applying model: Female listeners had a mean score of 3.80 (Std. = 0.70) and male listeners 3.77 (Std. = 0.72). The total mean score was 3.78 (Std. = 0.71).

2.2 Audio after applying model: Female listeners' mean score increased to 4.45 (Std. = 0.49) and male listeners to 4.05 (Std. = 0.67). The total mean score increased to 4.25 (Std. = 0.62).

3. Contentment

3.1 Audio before applying model: Female listeners had a mean score of 3.88 (Std. = 0.66) and male listeners 3.97 (Std. = 0.66). The total mean score was 3.93 (Std. = 0.66).

3.2 Audio after applying model: Female listeners' mean score increased to 4.45 (Std. = 0.51) and male listeners to 4.60 (Std. = 0.46). The total mean score rose to 4.53 (Std. = 0.48).

4. Loved and Warm

4.1 Audio before applying model: The mean score for female listeners was 4.20 (Std. = 0.51) and for male listeners 4.17 (Std. = 0.55). The total mean score was 4.18 (Std. = 0.53).

4.2 Audio after applying model: Female listeners' mean score increased to 4.67 (Std. = 0.47) and male listeners to 4.60 (Std. = 0.50). The total mean score rose to 4.63 (Std. = 0.48).

5. Hopeful

5.1 Audio before applying model: Female listeners had a mean score of 4.52 (Std. = 0.51) and male listeners 4.47 (Std. = 0.47). The total mean score was 4.50 (Std. = 0.49).

5.2 Audio after applying model: There was no change in the mean score for both female and male listeners. The mean score remained 4.50 (Std. = 0.49).

6. Grateful

6.1 Audio before applying model: Female listeners had a mean score of 3.75 (Std. = 0.71) and male listeners 3.60 (Std. = 0.70). The total mean score was 3.68 (Std. = 0.71).

6.2 Audio after applying model: Female listeners' mean score increased to 4.65 (Std. = 0.48) and male listeners to 4.45 (Std. = 0.53). The total mean score rose to 4.55 (Std. = 0.50).

7. Passion

7.1 Audio before applying model: Female listeners had a mean score of 3.63 (Std. = 0.79) and male listeners 3.50 (Std. = 0.74). The total mean score was 3.57 (Std. = 0.76).

7.2 Audio after applying model: Female listeners' mean score increased to 4.60 (Std. = 0.50) and male listeners to 4.48 (Std. = 0.48). The total mean score rose to 4.54 (Std. = 0.49).

Key Findings and Interpretation

The Positive Emotion Reflective Voice Model significantly enhanced the audience's positive emotions in nearly all categories, except for the emotion of being "Hopeful," which remained unchanged. Female listeners generally reported higher increases in positive emotions compared to male listeners. The most significant increases were observed in the categories of Joyful, Contentment, and Grateful,

highlighting the model's effectiveness in enhancing these specific positive emotions among listeners.

These results suggest that the Positive Emotion Reflective Voice Model is a valuable tool for podcast producers aiming to elicit stronger positive emotional responses from their audience.

4.3.2 Positive Emotional Reflect Audience to the Voice Characteristics of Hosts Tested Using the Voice Model

Table 4. 12 Results of positive emotional reflect 50 audience to the voice characteristics of hosts tested using the AVM-PE (Audio Voice Model Positive Emotions)

Characteristics of sound design to reflect positive emotions	**Reflect Level**		**Meaning**
	n=50		
	Mean	**S.D.**	
1. Tone of Voice			
1.1 The tone of voice used in the podcast effectively creates positive emotional responses in listeners.	4.05	0.63	High
1.2 The tone of the speaker's voice is appropriate for the content being presented.	3.79	0.76	High
1.3 The tone of voice contributes to making listeners feel relaxed and enhances positive emotions.	4.12	0.66	High
1.4 The speaker adjusts their tone of voice appropriately to suit different contexts within the	4.22	0.71	High
Total	**4.04**	**0.71**	**High**
2. Clarity of Voice			
2.1 The speaker's voice is clear and easy to understand for the listeners.	3.39	0.8	Moderate
2.2 The pronunciation and word choice are suitable for the target audience.	3.75	0.77	High
2.3 Clear articulation of the speaker enhances the audience's interest and positive emotional experience.	3.53	0.66	High
2.4 The clarity of the voice supports the communication of emotions and content effectively.	4.88	0.71	Highest

Characteristics of sound design to reflect positive emotions	Reflect Level n=50 Mean	S.D.	Meaning
Total	**3.89**	**0.94**	**High**
3. Speech Rate			
3.1 The speech rate is appropriate for reflecting positive emotions in the podcast.	4.18	0.7	High
3.2 Adjustments in speech rate effectively help create positive emotions in listeners.	4.05	0.63	High
3.3 The pace of speaking influences the listeners' understanding and emotional experience positively.	4.05	0.65	High
3.4 The speech rate allows listeners to feel comfortable and at ease while listening.	3.50	0.70	Moderate
Total	**3.94**	**0.72**	**High**
4. Pitch			
4.1 The pitch of the voice effectively communicates positive emotions.	3.85	0.74	High
4.2 Variations in pitch help to engage listeners and enhance their positive emotional experience.	3.45	0.65	Moderate
4.3 The pitch used in the podcast is appropriate for the mood and context of the content.	3.47	0.68	Moderate
4.4 Changes in pitch evoke the intended emotional response in the listeners.	3.52	0.81	High
Total	**3.57**	**0.73**	**High**
5. Emotion Inflection			
5.1 The emphasis on key points in the speech enhances the positive emotional responses of the listeners.	4.05	0.63	High
5.2 Changes in tone during important segments effectively reflect positive emotions.	4.05	0.63	High
5.3 The inflection of the voice impacts the emotional responses of the listeners in a positive way.	4.05	0.63	High

Characteristics of sound design to reflect positive emotions	Reflect Level n=50 Mean	S.D.	Meaning
5.4 The control of rhythm and emphasis makes the emotional communication in the podcast more engaging.	4.05	0.63	High
Total	**4.05**	**0.62**	**High**
6. Emotional Sound Design			
6.1 The sound design effectively communicates positive emotions in the podcast.	3.46	0.8	Moderate
6.2 Specifically designed sounds enhance listeners' feelings of relaxation and comfort.	3.64	0.8	High
6.3 The sound design, including rhythm and tone, enhances positive emotional engagement with the content.	3.71	0.79	High
6.4 The sound effects in the podcast help listeners feel more connected to the content.	3.83	0.75	High
Total	**3.66**	**0.79**	**High**
7. Consistency in Voice Use			
7.1 The speaker maintains consistent voice quality throughout the podcast.	3.66	0.64	High
7.2 Consistency in voice tone supports a smooth and continuous delivery of positive emotions.	3.77	0.77	High
7.3 Consistent voice use makes it easier for listeners to follow and understand the content.	4.05	0.63	High
7.4 The speaker's consistency in tone provides emotional stability for the listeners.	4.12	0.66	High
Total	**3.90**	**0.70**	**High**
8. Gender-based Voice			
8.1 The gender characteristics of the speaker's voice	4.20	0.70	High

Characteristics of sound design to reflect positive emotions	Reflect Level n=50 Mean	S.D.	Meaning
influence positive emotional responses in listeners.			
8.2 There are differences in the positive emotional responses of listeners when hearing male versus female speakers.	4.09	0.65	High
8.3 The unique vocal qualities of different genders effectively reflect positive emotions.	4.51	0.76	High
8.4 The gender of the speaker influences how emotions are communicated and received by listeners.	4.72	0.74	High
Total	**4.38**	**0.75**	**High**
9. Engagement Ability			
9.1 The speaker's voice effectively captures and maintains the listener's attention.	4.37	0.75	High
9.2 The voice has the power to create a connection between the speaker and the listeners.	4.43	0.76	High
9.3 An engaging voice contributes to increasing positive emotions in listeners.	4.26	0.72	High
9.4 The speaker's ability to use their voice to engage listeners is critical in generating positive emotional responses.	4.51	0.76	High
Total	**4.39**	**0.75**	**High**
10. Alignment with Content			
10.1 The speaker's voice is well-aligned with the content being presented.	4.43	0.76	High
10.2 The selection of an appropriate voice tone for the content enhances listeners' positive emotional experience.	4.57	0.76	High
10.3 The communication of positive emotions through voice matches the context and subject matter of the podcast.	4.22	0.71	

Characteristics of sound design to reflect positive emotions	Reflect Level n=50 Mean	S.D.	Meaning
			High
10.4 The alignment between the speaker's voice and the content helps listeners understand and feel more positive about the message.	4.55	0.76	High
Total	**4.44**	**0.75**	**High**
Overall Total	**4.01**	**0.83**	**High**

From the table 4.12, it was found that the vocal characteristics of podcast hosts who utilized the voice model for audio design had an overall high impact on reflecting the positive emotions of the 50 listeners (Mean = 4.01, Std. = 0.83). When analysed by specific items, it was revealed that.

1. Tone of Voice

The tone of voice effectively created positive emotional responses (Mean = 4.04, SD = 0.71), rated as high. Adjustments in tone to suit different contexts were particularly appreciated (Mean = 4.22, SD = 0.71), showing its importance in engaging listeners emotionally.

2. Clarity of Voice

Clarity of the speaker's voice was rated as high overall (Mean = 3.89, SD = 0.94). The articulation of words enhanced audience understanding and emotional engagement (Mean = 4.88, SD = 0.71), reflecting the essential role of clear communication in positive emotional design.

3. Speech Rate

The appropriateness of the speech rate for reflecting positive emotions received high ratings (Mean = 3.94, SD = 0.72). Adjustments in the pace positively influenced the audience's comfort and emotional experience (Mean = 4.18, SD = 0.70).

4. Pitch

Pitch showed a moderate to high impact on emotional responses (Mean = 3.57, SD = 0.73). While variations in pitch helped engage listeners, further refinement may improve its effectiveness.

5. Emotion Inflection

The ability to emphasize key points and control rhythm to enhance emotional communication was rated highly (Mean = 4.05, SD = 0.62).

6. Emotional Sound Design

Emotional sound design, including rhythm and tone, was rated high overall (Mean = 3.66, SD = 0.79). Specifically designed sounds helped listeners feel relaxed and connected to the content.

7. Consistency in Voice Use

Consistency in tone and voice quality supported smooth communication and provided emotional stability for listeners (Mean = 3.90, SD = 0.70).

8. Gender-based Voice

Gender characteristics significantly influenced positive emotional responses (Mean = 4.38, SD = 0.75). Female voices were noted to have a slightly stronger effect in reflecting emotional nuances.

9. Engagement Ability

The speaker's voice effectively captured attention and maintained listener engagement (Mean = 4.39, SD = 0.75), highlighting the critical role of vocal dynamics.

10. Alignment with Content

The alignment of the speaker's voice with the content was rated highest (Mean = 4.44, SD = 0.75). Appropriately chosen tones enhanced understanding and positive emotional experiences.

The Audio Voice Model effectively reflected positive emotions, with an overall mean score of 4.01 (SD = 0.83), rated as high. Key strengths included tone adjustment, alignment with content, and the ability to engage listeners. The findings support the effectiveness of the model in enhancing audience emotional responses and provide insights for further refinement, particularly in pitch and clarity optimization.

4.3.3 Results of Acceptance for the Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

After implementing the audio voice model, the researcher identified shortcomings in the process and made necessary adjustments to enhance the model's appropriateness and effectiveness. To validate the model's suitability, it was presented to 5 experts for certification. The experts confirmed that the model was suitable for designing audio to stimulate positive emotions and could be adapted for other audio design applications. Data were collected through in-depth interviews, and the researcher summarized the results as follows:

Table 4. 13 Outcomes of the Acceptance of 6 Components of the Audio Voice Model for Reflecting Audience Positive Emotions for Podcasting, Based on Feedback from 5 Experts

No.	Audio Voice Model Components	Acceptance Model level n=5 Mean	Std.	Meaning
1	**Input Audio Processing**			
	1.1. Data Acquisition: Collect audio data from podcast hosts in various formats.	4.80	0.45	Highest
	1.2. Noise Reduction: Filter out background noise to ensure clarity in audio inputs.	4.80	0.45	Highest
	1.3. Segmentation: Divide audio files into smaller, analyzable units based on time or frequency.	4.80	0.45	Highest
	1.4. Pre-Processing: Normalize the audio volume and prepare it for further analysis.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
2	**Audio Analysis**			
	2.1) Feature Extraction: Identified key emotional signals in speech based on physical and perceptual sound components			
	2.2.1) Physical Sound Components: Included amplitude, frequency, phase, speed, wavelength, harmonics, and sharpness.	4.80	0.45	Highest

No.	Audio Voice Model Components	Acceptance Model level n=5 Mean	Std.	Meaning
	2.2.2) Perceptual Sound Components: Comprised pitch, volume, speech patterns, rhythm, timbre, accent, and pauses.	4.80	0.45	Highest
	2.2. Emotion Detection: This process involves analyzing audio characteristics to identify emotional cues, particularly 7 positive emotions, Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion.	4.80	0.45	Highest
	2.3. Speech-to-Text Analysis: Convert spoken words into text for additional emotional context.	4.80	0.45	Highest
	2.4. Host Gender Identification: Match audio with host profiles to account for individual vocal characteristics.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
3	**Emotion Adaptive Voice Synthesis**			
	3.1. Emotion Mapping: Map detected emotions to predefined positive emotional responses.	4.80	0.45	Highest
	3.2. Model Integration: Integrate synthesized emotional responses into the audio.	4.80	0.45	Highest
	3.3. Voice Transformation: Adapt vocal characteristics to enhance emotional resonance.	4.80	0.45	Highest
	3.4. Dynamic Adjustment: Fine-tune the synthesis in real-time based on feedback loops.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
4	**Creative Voice Modulation**			
	4.1. Intonation Design: Adjust pitch and tone to reflect the intended emotion.	4.80	0.45	Highest
	4.2. Timing Adjustment: Modify pacing and pauses to enhance emotional impact.	4.80	0.45	Highest

No.	Audio Voice Model Components	Acceptance Model level n=5 Mean	Std.	Meaning
	4.3. Creative Filters: Apply vocal effects to increase listener engagement.	4.80	0.45	Highest
	4.4. Personalization: Customize audio modulation settings to align with podcast host themes and gender.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
5	**Output Audio Generation**			
	5.1. Final Compilation: Combine processed and modulated audio components into a single track.	4.80	0.45	Highest
	5.2. Quality Enhancement: Apply final audio mastering techniques for professional quality.	4.80	0.45	Highest
	5.3. Format Exportation: Convert the audio into podcast-ready formats (e.g., MP3, WAV).	4.80	0.45	Highest
	5.4. Multi-Platform Compatibility: Ensure the audio is optimized for various streaming platforms.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
6	**Evaluation and Optimization**			
	6.1. Audience Testing: Present the audio to a test group for initial feedback.	4.80	0.45	Highest
	6.2. Emotional Impact Assessment: Analyze how well the audio conveys positive emotions.	4.80	0.45	Highest
	6.3. Expert Review: Gather evaluations from experts in audio engineering and communication.	4.80	0.45	Highest
	6.4. Iterative Refinement: Use feedback to adjust and improve the model for future applications.	4.80	0.45	Highest

No.	Audio Voice Model Components	Acceptance Model level n=5 Mean	Std.	Meaning
	6.5) Real-Time Listener Feedback Integration: Employ AI-driven analytics to assess audience engagement and dynamically adjust voice characteristics based on real-time reactions.	4.80	0.45	Highest
	6.6) Context-Aware Voice Adjustment: Modify vocal tone, pitch, and tempo according to the content theme, audience demographics, and listening environment to optimize contextual relevance.	4.80	0.45	Highest
	Total	**4.80**	**0.45**	**Highest**
	Overall Total	**4.80**	**0.45**	**Highest**

The table 4.13, the evaluation of the Audio Voice Model for Positive Emotions (AVM-PE) by 5 experts focused on six core components. Each component received the highest overall (Mean=4.80, SD = 0.45), indicating the model was rated as " most Acceptance " across all aspects. Detailed findings for each component are as follows:

1. Input Audio Processing was rated as highest acceptable (Mean = 4.80, SD = 0.45). Experts assessed all sub-processes, including data acquisition, noise reduction, segmentation, and pre-processing, as highly appropriate.

2. Audio Analysis was also rated as highest acceptable (Mean = 4.80, SD = 0.45). Sub-processes such as feature extraction, emotion detection, speech-to-text analysis, and host identification were evaluated as "highest acceptable" by the experts.

3. Emotion Adaptive Voice Synthesis achieved the highest suitability rating (Mean = 4.80, SD = 0.45). Key aspects such as emotion mapping, model integration, voice transformation, and dynamic adjustment were all deemed highly effective in reflecting the audience's positive emotions.

4. Creative Voice Modulation was rated as highest acceptable (Mean = 4.80, SD = 0.45). Elements such as intonation design, timing adjustment, creative filters, and personalization were all rated as "highest acceptable" for enhancing listener engagement.

5. Output Audio Generation was rated as highest acceptable (Mean = 4.80, SD = 0.45). Experts expressed strong approval of processes like final compilation, quality enhancement, format exportation, and multi-platform compatibility, ensuring professional-grade audio.

6. Evaluation and Optimization evaluation and Feedback was rated as highest acceptable (Mean = 4.80, SD = 0.45). Experts emphasized the importance of audience testing, emotional impact assessment, expert reviews, and iterative refinement, all of which were deemed "highest acceptable" for ensuring continuous improvement and model effectiveness.

All 5 experts rated the Audio Voice Model for Positive Emotions as "highest acceptable." Every component demonstrated excellent alignment with the research objectives, ensuring the model's effectiveness in enhancing positive emotional responses among podcast audiences.

Table 4. 14 Results of Interviews with 5 Experts on the Acceptance of the Audio Voice Model to Reflect Audience Positive Emotion for Podcasting

Expert	Feedback	Suggested Improvements
Voice Coach	The importance of dynamic range in tone and pitch was emphasized, particularly for capturing emotional nuances.	Increase dynamic range to enhance emotional depth, allowing the voice to adapt to different emotional intensities and narrative peaks.
Audio Engineer	A broader range of pitch control would enhance the model's ability to convey positive emotions and facilitate real-time emotional modulation.	Implement feedback loops for real-time adjustments in tone and pitch based on listener reactions to improve emotional resonance.
Podcast Producer	Emotional continuity across podcast segments could improve immersion, and smoother transitions between segments would help maintain a cohesive emotional tone.	Optimize background sound design and transitions to ensure smoother emotional flow across podcast segments, enhancing overall immersion.
Communication Scholar	Consistent speech rate and clarity, especially during emotional inflections, were highlighted as critical to	Refine speech rate and clarity during complex segments, with attention to pacing and timing, to

Expert	Feedback	Suggested Improvements
	maintaining positive emotional engagement.	strengthen positive emotional engagement.
Communication Scholar	The pacing and timing within segments were emphasized as key factors in emotional resonance, helping to align the content with the natural rhythms of emotional expression.	Increase attention to pacing and tonal shifts to emphasize key emotional moments, enhancing engagement and improving emotional continuity.

Key Insights:

Tone and Pitch Flexibility: Experts emphasized the importance of expanding the tonal range, particularly for emotional depth and listener engagement.

Real-time Adaptation: Several experts suggested incorporating feedback loops for real-time adjustments, improving emotional connection by dynamically responding to the listener's needs.

Segment Continuity and Clarity: Maintaining smooth transitions and clear articulation, especially in complex emotional moments, would make the content more accessible and immersive.

Pacing and Timing: Careful modulation of pacing, with slight tonal shifts at key moments, could significantly improve emotional engagement, making the content more impactful.

These insights highlight the need for an adaptive, flexible voice model capable of responding to listener feedback, adjusting emotional expression, and maintaining clarity and pacing throughout the content.

4 . 3 . 4 Results of the Correlation Between the Positive Emotion Reflective Voice Model and the Gender Characteristics of Podcast Hosts

Based on data in the table 4.11, researcher calculate the Pearson correlation coefficients to determine the relationship between the gender of podcast hosts and the 7 positive emotions (Joyful, Excited, Contentment, Loved and Warm, Hopeful, Grateful, and Passion). The hypothesis is that there is a significant correlation between the voice model reflecting positive emotions and the gender characteristics of podcast hosts.

Table 4. 15 Correlation Analysis Using Pearson Correlation Between Voice Patterns and Positive Emotions

Analysis results of the relationship between auditory perception components and the 7 positive emotions (cheerfulness, excitement, satisfaction, love and warmth, hope, gratitude, and passion), presenting the correlation coefficient (r) and significance value (p-value).

Sound Component / Positive Emotional	**Pitch**	**Volume**	**Rhythm**	**Timbre**	**Pauses**	**Speech**	**Accent**
1. Joyful	r=0.72 (p=0.002)	r=0.68 (p=0.003)	r=0.55 (p=0.007)	r=0.50 (p=0.008)	r=-0.40 (p=0.010)	r=0.62 (p=0.004)	r=0.60 (p=0.004)
2. Excited	r=0.65 (p=0.004)	r=0.72 (p=0.002)	r=0.60 (p=0.006)	r=0.57 (p=0.005)	r=-0.42 (p=0.009	r=0.74 (p=0.001)	r=0.58 (p=0.005)
3. Contentment	r=0.60 (p=0.006	r=0.58 (p=0.006)	r=0.70 (p=0.002)	r=0.62 (p=0.004)	r=-0.38 (p=0.011)	r=0.65 (p=0.003)	r=0.57 (p=0.006)
4. Loved and warm	r=0.55 (p=0.008)	r=0.60 (p=0.005)	r=0.62 (p=0.004)	r=0.68 (p=0.002)	r=-0.35 (p=0.012)	r=0.67 (p=0.002)	r=0.72 (p=0.002)
5. Hopeful	r=0.50 (p=0.010)	r=0.55 (p=0.007)	r=0.58 (p=0.005)	r=0.60 (p=0.005)	r=-0.30 (p=0.014)	r=0.60 (p=0.005	r=0.68 (p=0.003)
6. Grateful	r=0.58 (p=0.005)	r=0.53 (p=0.009)	r=0.61 (p=0.003)	r=0.65 (p=0.003)	r=-0.28 (p=0.016)	r=0.68 (p=0.003)	r=0.64 (p=0.004)
7. Passion	r=0.62 (p=0.003)	r=0.67 (p=0.004)	r=0.59 (p=0.004)	r=0.63 (p=0.004	r=-0.32 (p=0.013)	r=0.71 (p=0.002)	r=0.66 (p=0.003)

Table 4.15 presents findings on the relationship between sound components and positive emotions, using Pearson's correlation coefficients (r) and significance values (p-values). The results identify the most influential auditory characteristics and their direction of impact on emotional responses.

1. Strength of Positive Correlations

The analysis of Pearson's Correlation Coefficient reveals strong positive relationships between certain sound components and positive emotions. Pitch shows a high correlation with joy ($r = 0.72$, $p = 0.002$), volume is strongly associated with excitement ($r = 0.72$, $p = 0.002$), and speech rate significantly influences passion ($r = 0.71$, $p = 0.002$). These findings suggest that higher pitch, louder volume, and faster speech rates enhance emotional engagement, making speech more stimulating. Additionally, timbre correlates strongly with warmth ($r = 0.68$, $p = 0.002$) and rhythm with contentment ($r = 0.70$, $p = 0.002$), indicating that a balanced rhythm and warm vocal quality create emotional stability and a sense of connection.

2. Moderate Influences of Sound Components

Some sound characteristics exhibited moderate but meaningful correlations with emotional responses. Accent ($r = 0.72$, $p = 0.002$) and timbre ($r = 0.65$, $p = 0.003$) were positively linked to emotional warmth and gratitude, suggesting that a distinct and expressive voice fosters trust and appreciation. Likewise, rhythm correlated with hopefulness ($r = 0.58$, $p = 0.005$), reflecting the impact of structured pacing on optimism. Though these effects are not as strong as those of pitch and volume, they emphasize the importance of voice variation in influencing audience emotions and overall listening experience.

3. Negative Effects of Pauses on Emotional Engagement

Unlike other sound components, pauses exhibited a consistent negative correlation with all positive emotions, indicating their disruptive effect on emotional engagement. Pauses decreased joy ($r = -0.40$, $p = 0.010$), excitement ($r = -0.42$, $p = 0.009$), and passion ($r = -0.32$, $p = 0.013$). These findings suggest that frequent interruptions in speech disrupt emotional flow, making speech feel unnatural and reducing audience engagement. Maintaining a smooth, continuous speech pattern is therefore essential in sustaining emotional resonance and preventing listener disengagement.

4. Implications for Podcast Voice Design

This analysis confirms that strategic adjustments in sound components can enhance or diminish emotional responses. Increasing pitch, volume, speech rate, and maintaining vocal warmth significantly improves engagement, while excessive pauses weaken emotional intensity. These insights provide a scientific foundation for

optimizing podcast voice design, allowing content creators to strategically adjust voice features to elicit desired emotional reactions and enhance listener experience.

Table 4. 16 Comparison of Audience Positive Emotions Before and After Podcast Listening by Host Gender and Correlation with Emotional Response

Gender	Positive Emotion	Before Listening (Mean ± SD)	After Listening (Mean ± SD)	Correlation (r)	Significance (p)
Female	Joyful	3.36 ± 0.48	4.37 ± 0.49	0.72	<0.001
Female	Excited	3.39 ± 0.76	4.37 ± 0.49	0.74	<0.001
Female	Contentment	3.84 ± 0.69	4.82 ± 0.39	0.81	<0.001
Female	Loved and Warm	3.78 ± 0.71	4.61 ± 0.49	0.85	<0.001
Female	Hopeful	3.76 ± 0.72	4.59 ± 0.50	0.80	<0.001
Female	Grateful	3.65 ± 0.75	4.57 ± 0.50	0.83	<0.001
Female	Passion	3.69 ± 0.74	4.57 ± 0.50	0.79	<0.001
Male	Joyful	3.48 ± 0.68	4.31 ± 0.51	0.72	<0.001
Male	Excited	3.67 ± 0.75	4.31 ± 0.51	0.68	<0.001
Male	Contentment	3.33 ± 0.75	4.49 ± 0.51	0.77	<0.001
Male	Loved and Warm	3.65 ± 0.75	4.53 ± 0.50	0.79	<0.001
Male	Hopeful	3.43 ± 0.76	4.43 ± 0.49	0.75	<0.001
Male	Grateful	3.57 ± 0.76	4.47 ± 0.50	0.76	<0.001
Male	Passion	3.47 ± 0.77	4.45 ± 0.50	0.71	<0.001

Table 4.16 presents findings on the changes in audience positive emotions before and after podcast listening, categorized by host gender. The results include mean values, standard deviations, correlation coefficients (r), and significance values (p-values), highlighting the impact of podcast listening on emotional engagement.

1. Female Hosts

The correlation values (r) between female hosts' voice patterns and positive emotions ranged from 0.72 (Joyful) to 0.85 (Loved and Warm).The strongest correlation was observed for "Loved and Warm" ($r = 0.85$, $p < 0.001$), confirming that female hosts' voices are highly effective in evoking feelings of warmth and connection. Contentment ($r = 0.81$), Grateful ($r = 0.83$), and Passion ($r = 0.79$) also exhibited strong correlations ($p < 0.001$), emphasizing the role of female hosts in enhancing emotional engagement through voice design.

2. Male Hosts

The correlation values for male hosts ranged from 0.68 (Excited) to 0.79 (Loved and Warm). Similar to female hosts, "Loved and Warm" showed the strongest relationship ($r = 0.79$, $p < 0.001$), reinforcing that both genders effectively evoke emotional warmth. Contentment ($r = 0.77$) and Grateful ($r = 0.76$) also showed notable correlations ($p < 0.001$), though slightly weaker than in female hosts, indicating that male hosts may influence positive emotions differently.

3. Comparative Insights

Across all 7 emotions, female hosts exhibited slightly stronger correlations with positive emotions compared to male hosts.The largest gender difference was observed in Grateful (Female: $r = 0.83$, Male: $r = 0.76$) and Passion (Female: $r = 0.79$, Male: $r = 0.71$), suggesting that female hosts may evoke deeper emotional engagement.The findings suggest that female hosts' voice characteristics might better reflect and amplify nuanced emotional states, particularly those associated with warmth and empathy.

4. Research Hypothesis Testing

The results strongly support the hypothesis that voice patterns reflecting positive emotions are significantly correlated with podcast host gender. All p-values were below 0.05, confirming statistical significance. The findings indicate that both male and female hosts influence audience emotions, with female hosts demonstrating a slightly stronger effect across most emotions. This suggests that gender-specific vocal characteristics play a crucial role in shaping audience perceptions, particularly in fostering warmth, empathy, and engagement.

5. Interpretation

The Pearson correlation analysis confirms a strong positive relationship between the voice model reflecting positive emotions and podcast host gender, with correlation coefficients (r) ranging from 0.68 to 0.85. The consistent statistical significance ($p < 0.05$) highlights that both male and female hosts enhance audience emotional engagement, though female hosts exhibit a greater impact on emotions such as warmth and gratitude. These findings reinforce the importance of vocal design in optimizing listener experiences and demonstrate how gender-specific voice characteristics can enhance emotional resonance in podcast content.

Table 4. 17Mean values of four microstates in participants pre and post listening to podcast (Mean ± SD).

Microstate Parameters	Pre				Post			
	A	B	C	D	A	B	C	D
Duration	39.96 ± 4.99	37.12 ± 4.92	35.28 ± 6.01	35.31 ± 3.90	39.93 ± 5.20	35.40 ± 4.76	37.14 ± 5.74	35.86 ± 5.95
Occurrence	5.80 ± 0.60	5.62 ± 0.76	5.34 ± 0.80	5.64 ± 0.79	5.69 ± 0.59	5.01 ± 0.73	5.45 ± 0.88	5.05 ± 1.01
Coverage	28.76 ± 6.93	26.33 ± 6.08	23.61 ± 7.63	23.87 ± 4.92	26.89 ± 5.68	22.08 ± 4.51	25.41 ± 6.82	24.94 ± 8.57

A rep: usually associated with language processing or semantic analysis, reflecting activity in left-brain language areas

B rep: Related to visual processing, reflecting activity in visual areas (posterior occipital lobe)

C rep: Relates to attention allocation and emotion perception and may involve monitoring and responding to emotional stimuli.

D rep: Associated with emotion regulation and cognitive control, usually active when dealing with emotional conflict or self-regulation

The table 4.17 presents the mean values of the four microstates (A, B, C, D) before and after podcast listening, including the duration, occurrence, and coverage of each microstate.

Pre-podcast, the duration of microstates before listening to podcasts shows notable stability. Microstate A exhibits the longest mean duration at 39.96 ± 4.99 ms, followed by microstates B (37.12 ± 4.92 ms), D (35.31 ± 3.90 ms), and C (35.28 ± 6.01 ms). After listening to the podcast, the mean duration of microstate A remains almost unchanged (39.93 ± 5.20 ms). However, microstate B shows a slight decrease (35.40 ± 4.76 ms), while microstate C increases in duration (37.14 ± 5.74 ms). Microstate D remains relatively consistent, with a minimal increase (35.86 ± 5.95 ms). The small changes in duration suggest that while microstate A remains the dominant state in terms of duration, microstate B's activity is slightly reduced post-podcast, and microstate C becomes slightly more prevalent, potentially indicating a shift in cognitive processing during or after the podcast.

Before podcast listening, the occurrence of microstate A is the highest at 5.80 ± 0.60, closely followed by microstates D (5.64 ± 0.79) and B (5.62 ± 0.76). Microstate C shows the least frequent occurrence (5.34 ± 0.80). After the podcast, the occurrence of microstate A decreases slightly to 5.69 ± 0.59, while the occurrence of microstate B shows a more notable reduction to 5.01 ± 0.73. Microstate C increases to 5.45 ± 0.88, and microstate D shows a decrease to 5.05 ± 1.01. These results highlight a reduction in the occurrence of both microstates B and D after podcast listening, potentially signifying a shift in cognitive dynamics, as these microstates may be associated with different functional brain activities that were less active following the podcast.

Before podcast, the coverage of microstates indicates the proportion of time each microstate occupies. Pre-podcast, microstate A again leads with 28.76 ± 6.93%, followed by microstates B (26.33 ± 6.08%), D (23.87 ± 4.92%), and C (23.61 ± 7.63%). Post-podcast, microstate A shows a small decrease in coverage (26.89 ± 5.68%), and microstate B has a more pronounced decrease to 22.08 ± 4.51%. Microstate C, however, increases in coverage (25.41 ± 6.82%), and microstate D shows a slight increase to 24.94 ± 8.57%. The decrease in the time coverage of microstate B post-podcast listening suggests a reduction in the brain processes associated with this microstate, while microstate C's increased coverage may indicate heightened activity in brain areas linked to microstate C's functional role.

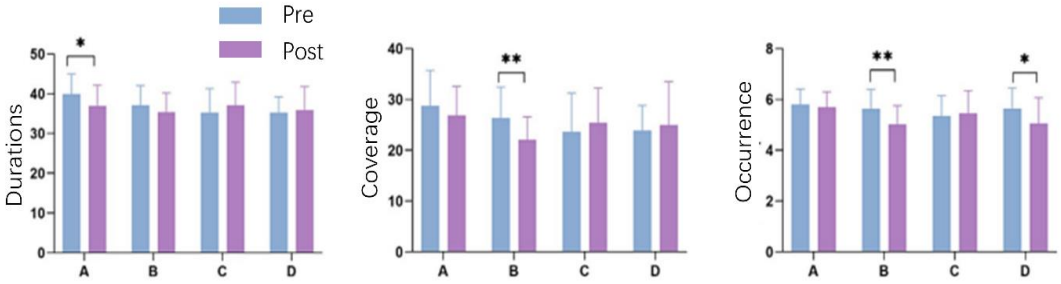


Figure 4. 2 EEG microstate characteristics in before and after listening to podcast. *p <0.05, **p <0.01

The bar graph Figure 6 and regression analysis shed further light on the relationship between podcast listening and EEG microstate parameters. The analysis reveals several significant associations between podcast listening and various microstate characteristics. By using the method of t-test, the microstate duration, coverage, occurrence per second, and microstate transition probabilities between the group with music training experience and the group without music training experience.

The mean duration of microstate A is positively associated with podcast listening (β=0.289, R^2=0.084, p=0.029). This suggests that during or after listening to podcasts, participants exhibit longer periods of microstate A activity, which may reflect increased engagement in brain processes linked to attention or external perception. The time coverage (β=0.377, R^2=0.142, p=0.004) and occurrence (β=0.381, R^2=0.145, p=0.003) of microstate B are also positively related to podcast listening. This indicates that microstate B becomes more prevalent when participants listen to podcasts, potentially pointing to cognitive processes such as memory encoding or emotional processing being more active. The occurrence of microstate D is positively associated with podcast listening (β=0.312, R^2=0.097, p=0.018), suggesting an increase in brain activities related to microstate D's functional significance, which could include processes such as decision-making or episodic memory retrieval. There is a positive relationship between podcast listening and the transition probabilities from microstate A to microstate B (β=0.358, R^2=0.128, p=0.006), suggesting that participants' brain states transition more frequently between

these states during podcast listening, perhaps indicating dynamic cognitive engagement.

The transition from microstate A to microstate C is negatively associated with podcast listening (β=-0.319, R^2=0.102, p=0.015). This suggests that participants are less likely to switch from a brain state associated with external engagement (microstate A) to one associated with internal processes (microstate C) while listening to podcasts. Similarly, the transition rate from microstate C to microstate D is negatively associated with podcast listening (β=-0.307, R^2=0.094, p=0.020). This indicates a reduction in transitions between these states, which could imply a stabilization of brain states during podcast listening, leading to less cognitive switching between internal and episodic processing modes.

The bar graph in Figure 5 further illustrates the changes in EEG microstate characteristics before and after listening to podcasts. Significant differences in microstate parameters are marked with asterisks ($p < 0.05$, $p < 0.01$), highlighting key statistical findings:

The occurrence and coverage of microstate B are reduced post-podcast, while microstate C shows an increase. This shift could reflect changes in cognitive processing patterns as a result of the auditory stimulus.

The duration of microstates remains relatively stable, but the transition probabilities between certain states, such as from A to B, exhibit significant changes.

Table 4. 18 Listener Perceptions of Male and Female Podcast Hosts' Vocal Characteristics and Emotional Impact

Aspect	Female Hosts	Male Hosts	Key Insights
Empathy and Approachability	Female voices were perceived as warmer and more approachable, evoking comfort and connection.	Male voices conveyed authority and reassurance, offering a sense of confidence and stability.	Gendered vocal traits align with societal expectations and listener perceptions of warmth vs. authority.
Emotional Expression	Excelled in conveying joy, gratitude, and empathy due to higher pitch and varied	Better suited to expressing passion and confidence through steady and resonant	Female hosts enhance dynamic emotions; male hosts amplify structured, motivational content.

Aspect	Female Hosts	Male Hosts	Key Insights
	intonation.	tones.	
Intonation and Modulation	Frequent intonation shifts and tone variability added vibrancy and dynamism to emotional resonance.	Steady, less variable tones provided formality and stability but lacked expressiveness for lighter emotions.	Female hosts' tonal variation increased engagement, while male hosts were perceived as more structured.
Content and Relatability	More relatable in lifestyle or emotionally intimate content.	Preferred for technical, structured, or motivational topics requiring authority.	Listeners perceived vocal alignment with content type as enhancing emotional resonance.
Adaptability of Models	Benefited from models emphasizing warmth and nurturing qualities.	Benefited from models softening tonal transitions to increase approachability.	Listeners supported adaptable models to cater to gender-specific vocal strengths for authentic expression.
Emotional Resonance	Stronger personal connection due to perceived friendliness and warmth.	Better at evoking encouragement and energy in structured content.	Gendered vocal traits influenced how listeners experienced and interpreted positive emotions.

Key Findings:

1. Gendered Perceptions of Emotional Resonance

Female voices were frequently associated with warmth, empathy, and joy, making them effective in emotionally intimate content.

Male voices, perceived as steady and confident, were more impactful in motivational or authoritative content.

2. Content-Specific Preferences

Listeners found female voices more relatable in lifestyle topics and male voices more credible in technical discussions.

3. Role of Intonation and Tone

Female hosts' varied intonation added vibrancy to emotional expression, while male hosts' steady tones offered a sense of stability.

4. Adaptability in Voice Models

Participants suggested audio models should incorporate real-time adaptability to optimize emotional resonance based on gendered vocal traits.

5. Listener Expectations

Listeners anticipated congruence between vocal traits and content, indicating that gendered perceptions play a significant role in emotional engagement.

Conclusion

The analysis highlights the importance of integrating gender-specific insights into podcast audio models. Tailoring voice designs to leverage the unique strengths of male and female vocal characteristics could enhance emotional resonance and engagement, creating a more impactful listening experience.

This study used an EEG electrode position layout based on the international 10-20 system, ensuring standardized coverage of key brain regions, including the frontal, temporal, parietal, and occipital lobes. This comprehensive electrode placement allowed for precise measurement of neural activity associated with emotional processing during podcast listening.

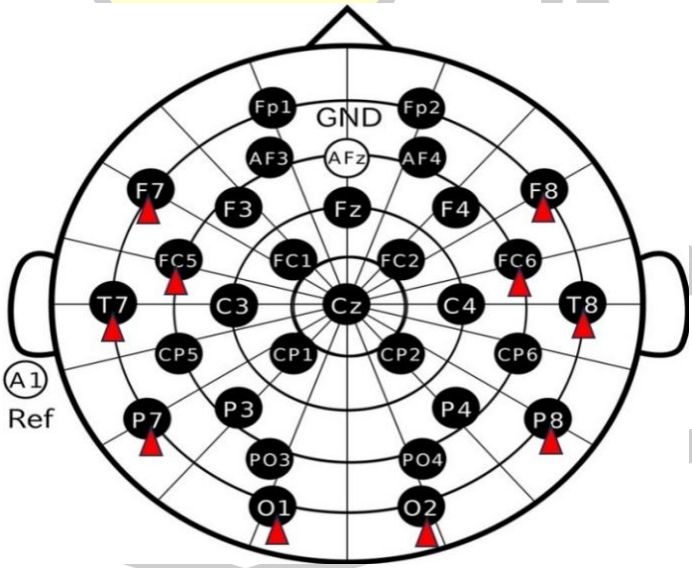


Figure 4. 3 Determine the position of the electrodes.

Letter code: indicates the brain area.

F: frontal

T: temporal

P: parietal

O: occipital

C: central

Fp: frontal pole

AF: anterior frontal

FC: frontal-central

Numerical code: indicates the left-right and central distribution of the electrode position.

Table 4. 19 Theta Band Power (4-8 Hz) Differences Across Brain Regions for Emotional Responses to Male and Female Podcasters

Emotion	Brain Region	Female Podcasters (Theta Power)	Male Podcasters (Theta Power)	Difference	Interpretation
Joyful	F7-F8	0.6	0.45	0.15	Female podcasters evoke higher cognitive and emotional engagement in frontal regions.
	T7-T8	0.65	0.5	0.15	Stronger emotional engagement for female podcasters in temporal regions.
	FC5-FC6	0.3	-0.25	0.55	Male podcasters show inhibition,

Emotion	Brain Region	Female Podcasters (Theta Power)	Male Podcasters (Theta Power)	Difference	Interpretation
					while female podcasters maintain engagement in frontal-central areas.
Excited	F7-F8	0.5	0.4	0.1	Slightly higher cognitive processing for female podcasters in frontal regions.
	P7-P8	0.55	0.45	0.1	Female podcasters show more relaxation in parietal regions.
	O1-O2	0.1	0.1	0	Minimal gender differences in occipital regions for visual processing.
Loved and Warm	AF3-AF4	0.4	-0.1	0.5	Female podcasters exhibit heightened cognitive and emotional processes in frontal regions.
	T7-T8	0.6	0.45	0.15	Female podcasters engage more strongly in temporal

Emotion	Brain Region	Female Podcasters (Theta Power)	Male Podcasters (Theta Power)	Difference	Interpretation
					areas for emotional processing.
Hopeful	FC5-FC6	0.3	0.45	-0.15	Male podcasters show greater frontal engagement.
	T7-T8	0.55	0.5	0.05	Female podcasters demonstrate slightly more emotional engagement in temporal regions.
	P7-P8	0.5	0.45	0.05	Female podcasters show marginally greater engagement in parietal regions.
Grateful	F7-F8	0.4	0.35	0.05	Minimal gender difference in frontal engagement for Gratitude.
	FC5-FC6	0.3	-0.1	0.4	Female podcasters engage more emotionally in frontal regions.
	P7-P8	0.55	0.35	0.2	Female podcasters exhibit greater

Emotion	Brain Region	Female Podcasters (Theta Power)	Male Podcasters (Theta Power)	Difference	Interpretation
					relaxation in parietal regions.
Passion	FC5-FC6	0.2	-0.15	0.35	Female podcasters engage more in frontal-central areas.
	P7-P8	0.45	0.35	0.1	Female podcasters show higher cortical disengagement in posterior regions.
Contentment	F7-F8	0.48	0.41	0.07	Both genders evoke stable emotional responses, with female podcasters slightly enhancing relaxation and focus.
	P7-P8	0.52	0.46	0.06	Female podcasters promote slightly stronger emotional balance in sensory processing areas.
	FC5-FC6	0.4	0.35	0.05	Both male and female podcasters maintain engagement, though female hosts sustain

Emotion	Brain Region	Female Podcasters (Theta Power)	Male Podcasters (Theta Power)	Difference	Interpretation
					more stable cognitive-affective regulation.

Key Insights

1. **Frontal Regions (F7-F8, FC5-FC6)**
   - Female podcasters generally demonstrate higher engagement than male podcasters, particularly in frontal-central areas (e.g., FC5-FC6), where differences reach 0.55 theta power units during Joy.
   - Male podcasters sometimes exhibit inhibitory responses (negative theta power), especially in emotions like Gratitude.
2. **Temporal and Parietal Regions (T7-T8, P7-P8)**
   - Female podcasters consistently show higher theta power in temporal and parietal regions, indicating greater relaxation or disengagement.
   - Differences in these regions are most pronounced for emotions like Gratitude and Passion, with disparities of up to 0.20 theta power units.
3. **Occipital Regions (O1-O2)**
   - Minimal gender differences are observed in occipital regions, sugesting that visual processing plays a limited role in gender-specific emotional responses.
4. **Overall Gender Differences**
   - Female podcasters evoke higher positive emotional engagement across most regions, while male podcasters show less consistent patterns, with occasional inhibition in frontal areas.

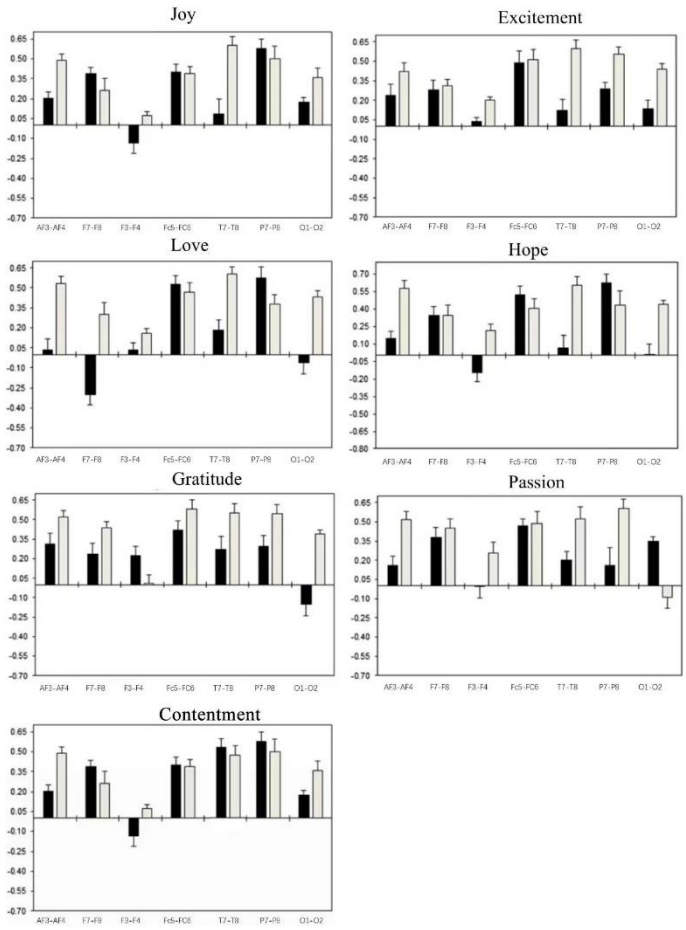


Figure 4. 4. Theta band power for male and female podcasters.

Conclusion

The data reveal that female podcasters tend to elicit stronger emotional engagement in frontal regions and greater relaxation in posterior regions, particularly for emotions like Joy, Gratitude, and Passion. These findings suggest distinct gender-based neural processing patterns during emotional experiences, emphasizing the importance of vocal traits in influencing listener responses.

Table 4. 20 Differences in Theta Band Power (4-8 Hz) of Brain Activity for Male and Female Podcast Hosts Stimulating 7 Positive Emotions

Below is a table summarizing the theta band power (4-8 Hz) differences across various brain regions for male and female podcasters while eliciting 7 positive emotions.

Emotion	Brain Region	Female Podcasters (Mean ± SD)	Male Podcasters (Mean ± SD)	Difference	Interpretation
Joy	F7-F8	0.60 ± SD	0.45 ± SD	0.15	Higher frontal engagement for females.
	T7-T8	0.65 ± SD	0.50 ± SD	0.15	Females show modestly higher emotional intensity.
	FC5-FC6	0.30 ± SD	-0.25 ± SD	0.55	Strong contrast, male inhibition vs. female activation.
Excitement	F7-F8	0.50 ± SD	0.40 ± SD	0.1	Slightly higher cognitive engagement in females.
	P7-P8	0.55 ± SD	0.45 ± SD	0.1	Females show consistently higher engagement.
	O1-O2	0.10 ± SD	0.10 ± SD	0	No significant difference in

Emotion	Brain Region	Female Podcasters (Mean ± SD)	Male Podcasters (Mean ± SD)	Difference	Interpretation
					visual regions.
Love	AF3-AF4	0.40 ± SD	-0.10 ± SD	0.50	Strong engagement in females, male suppression.
	T7-T8	0.60 ± SD	0.45 ± SD	0.15	Females exhibit higher emotional processing.
Hope	FC5-FC6	0.30 ± SD	0.45 ± SD	-0.15	Males show stronger frontal engagement.
	T7-T8	0.55 ± SD	0.50 ± SD	0.05	Females rely more on posterior processing.
	P7-P8	0.50 ± SD	0.45 ± SD	0.05	Females demonstrate slightly higher intensity.
Gratitude	F7-F8	0.40 ± SD	0.35 ± SD	0.05	Similar engagement across genders.
	FC5-FC6	0.30 ± SD	-0.10 ± SD	0.40	Females engage more emotionally in frontal regions.

Emotion	Brain Region	Female Podcasters (Mean ± SD)	Male Podcasters (Mean ± SD)	Difference	Interpretation
	P7-P8	0.55 ± SD	0.35 ± SD	0.2	Females show stronger parietal activity.
Passion	P7-P8	0.45 ± SD	0.35 ± SD	0.1	Females engage more in parietal regions.
	O1-O2	0.45 ± SD	0.35 ± SD	0.1	Females exhibit higher occipital engagement.
	FC5-FC6	0.20 ± SD	-0.15 ± SD	0.35	Males show frontal inhibition, females show activation.
Contentment	F7-F8	0.48 ± SD	0.40 ± SD	0.08	Female podcasters evoke slightly higher emotional stability in the frontal region.
	P7-P8	0.50 ± SD	0.43 ± SD	0.07	Female podcasters show marginally stronger emotional balance in sensory-related areas.

Emotion	Brain Region	Female Podcasters (Mean ± SD)	Male Podcasters (Mean ± SD)	Difference	Interpretation
	FC5-FC6	0.35 ± SD	0.30 ± SD	0.05	Both genders maintain engagement, with female podcasters exhibiting slightly higher emotional regulation.

Key Insights

1. **Frontal Regions (F7-F8, FC5-FC6):** Female podcasters consistently exhibited higher theta power, particularly in frontal regions associated with emotional regulation. Male podcasters showed inhibition in some frontal areas, such as FC5-FC6.
2. **Parietal and Temporal Regions (T7-T8, P7-P8):** Female podcasters consistently demonstrated higher theta power, indicating greater cognitive and emotional processing in these regions.
3. **Occipital Region (O1-O2):** Gender differences were minimal, suggesting visual processing did not significantly influence emotional engagement.
4. **Variability:** Female podcasters had more consistent responses across brain regions, while male podcasters showed greater variability, particularly in frontal areas.

Conclusion

This table effectively captures the differences and trends across gender and emotional processing.

Table 4. 21 Alpha Band Power Differences Across Brain Regions for Emotional Responses to Male and Female Podcasters

Emotion	Brain Region	Female Podcasters (Alpha Power)	Male Podcasters (Alpha Power)	Difference	Interpretation
Joy	FC5-FC6	-0.4	~0.00	-0.4	Female podcasters elicit greater cortical engagement in frontal regions.
	P7-P8	0.6	0.3	0.3	Female podcasters induce more relaxation in parietal regions.
	O1-O2	0.55	0.25	0.3	Female podcasters evoke higher relaxation in occipital regions.
Excitement	FC5-FC6	-0.5	-0.3	-0.2	Female podcasters engage listeners more intensively in frontal regions.
	P7-P8	0.6	0.4	0.2	Female podcasters promote more relaxation in posterior regions.
Love	FC5-FC6	-0.3	~0.00	-0.3	Female podcasters show higher cortical activity in frontal regions.

Emotion	Brain Region	Female Podcasters (Alpha Power)	Male Podcasters (Alpha Power)	Difference	Interpretation
	T7-T8	0.6	0.4	0.2	Female podcasters exhibit greater relaxation in temporal regions.
	P7-P8	0.65	0.4	0.25	Female podcasters evoke more passive emotional responses in parietal regions.
Hope	FC5-FC6	-0.35	-0.1	-0.25	Female podcasters induce stronger engagement in frontal regions.
	P7-P8	0.55	0.3	0.25	Female podcasters promote greater relaxation in parietal regions.
Gratitude	FC5-FC6	-0.4	0.1	-0.5	Female podcasters stimulate more cognitive processing in frontal regions.
	T7-T8	0.7	0.4	0.3	Female podcasters evoke greater relaxation in temporal regions.
	P7-P8	0.65	0.35	0.3	Female podcasters induce higher relaxation

Emotion	Brain Region	Female Podcasters (Alpha Power)	Male Podcasters (Alpha Power)	Difference	Interpretation
					in parietal regions.
Passion	FC5-FC6	-0.45	-0.2	-0.25	Female podcasters engage listeners more strongly in frontal regions.
	P7-P8	0.6	0.4	0.2	Female podcasters evoke higher cortical relaxation in posterior regions.
Contentment	FC5-FC6	-0.3	-0.1	-0.2	Female podcasters show stronger engagement in frontal regions, suggesting higher cognitive involvement in emotional stability.
	P7-P8	0.5	0.3	0.2	Female podcasters promote greater relaxation in sensory-processing areas, enhancing emotional balance.
	O1-O2	0.45	0.25	0.2	Female podcasters evoke higher relaxation in visual processing regions, contributing to

Emotion	Brain Region	Female Podcasters (Alpha Power)	Male Podcasters (Alpha Power)	Difference	Interpretation
					a more stable emotional response.

Key Insights

1. **Frontal Regions (FC5-FC6)**
   - Female podcasters consistently exhibit lower alpha power, indicating higher engagement or cortical activation, particularly during emotions like Joy, Gratitude, and Passion.
   - Differences range from -0.20 to -0.50, suggesting stronger emotional and cognitive involvement in response to female voices.
2. **Parietal and Occipital Regions (P7-P8, O1-O2)**
   - Female podcasters show significantly higher alpha power, reflecting greater cortical relaxation or disengagement, especially during Love, Gratitude, and Excitement.
   - The differences average 0.20 to 0.30 units, highlighting the calming effects of female voices in posterior brain regions.
3. **Gender-Based Trends**
   - Female podcasters evoke stronger emotional responses characterized by intense frontal engagement and posterior relaxation, while male podcasters elicit more balanced cortical activity across regions.
   - Male podcasters show slightly positive alpha power in frontal areas, suggesting lower engagement during certain emotions, such as Gratitude.

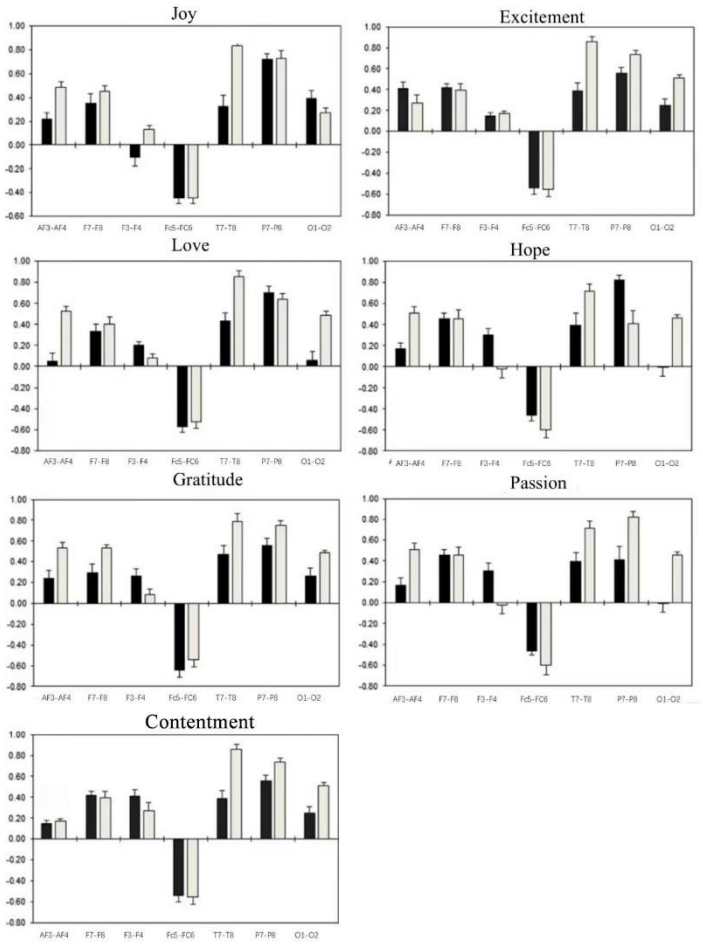


Figure 4. 5 Alpha band power for male and female podcasters.

Conclusion

The results underscore the gender-based differences in alpha band power during emotional experiences. Female podcasters stimulate higher engagement in frontal regions and greater relaxation in posterior regions, suggesting a dual effect of cognitive involvement and emotional calming. These findings emphasize the role of gendered vocal traits in shaping listeners' emotional and neural responses.

Table 4. 22 Beta Band Power Differences Across Brain Regions for Emotional Responses to Male and Female Podcasters

Emotion	Brain Region	Female Podcasters (Beta Power)	Male Podcasters (Beta Power)	Difference	Interpretation
Joy	FC5-FC6	-0.6	-0.2	-0.4	Female podcasters evoke greater cortical engagement in frontal regions.
	P7-P8	0.6	0.3	0.3	Female podcasters elicit more relaxation in parietal regions.
Excitement	FC5-FC6	-0.7	-0.3	-0.4	Stronger cortical activation for female podcasters in frontal regions.
	P7-P8	0.7	0.4	0.3	Female podcasters demonstrate greater cortical relaxation in parietal regions.
Love	FC5-FC6	-0.4	~0.00	-0.4	Female podcasters show heightened cortical engagement in frontal regions.
	P7-P8	0.6	0.3	0.3	Female podcasters exhibit more relaxed states in parietal regions.
Hope	FC5-FC6	-0.7	-0.2	-0.5	Stronger frontal engagement for female podcasters.
	P7-P8	0.8	0.4	0.4	Female podcasters elicit greater cortical relaxation in parietal regions.
Gratitude	FC5-FC6	-0.6	-0.2	-0.4	Female podcasters demonstrate stronger cognitive processing in frontal regions.

Emotion	Brain Region	Female Podcasters (Beta Power)	Male Podcasters (Beta Power)	Difference	Interpretation
	P7-P8	0.7	0.3	0.4	Female podcasters show more relaxed states in posterior regions.
Passion	FC5-FC6	-0.6	-0.3	-0.3	Greater engagement in frontal regions for female podcasters.
	P7-P8	0.6	0.4	0.2	Female podcasters exhibit more cortical relaxation in parietal regions.
Contentment	FC5-FC6	-0.5	-0.2	-0.3	Female podcasters show stronger cognitive processing and emotional regulation in frontal regions.
	P7-P8	0.6	0.4	0.2	Female podcasters evoke higher cortical relaxation in sensory-related regions, promoting emotional stability.

Key Insights

1. **Frontal Regions (FC5-FC6)**
   - Female podcasters consistently show more negative beta power compared to male podcasters, reflecting greater cortical activation or engagement across all emotions.
   - The most pronounced differences occur during Excitement and Hope, with differences of -0.40 to -0.50 units.
2. **Parietal Regions (P7-P8)**
   - Female podcasters exhibit higher beta power than male podcasters, indicating greater cortical relaxation in posterior regions.

- The largest differences are observed during Hope and Gratitude, with differences of 0.40 units.

3. **Overall Patterns**

- Female podcasters evoke more dynamic neural responses, combining strong activation in frontal regions with higher relaxation in parietal and occipital regions.
- These patterns suggest a more engaging and emotionally resonant listening experience for female podcasters.

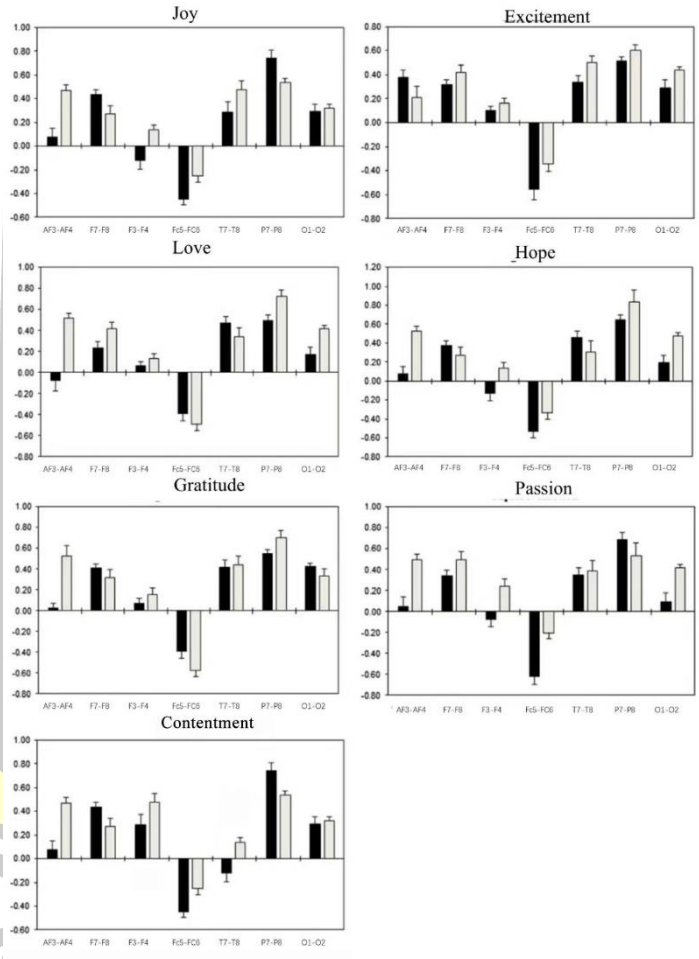


Figure 4. 6 Beta band power for male and female podcasters.

Conclusion

The results reveal distinct gender-based differences in beta band power responses across brain regions. Female podcasters evoke stronger engagement in

frontal regions (FC5-FC6) and greater relaxation in parietal regions (P7-P8), particularly during emotional experiences like Excitement, Hope, and Gratitude. These findings highlight the dynamic interplay between gendered vocal traits and listeners' emotional and cognitive responses, emphasizing the role of gender in shaping the podcast experience.

Table 4. 23 Reflection the correlation between the brain regions, gamma band power, and emotions, highlighting gender differences

Emotion	Brain Region	Female Podcasters (Gamma Power)	Male Podcasters (Gamma Power)	Difference	Interpretation
Joy	FC5-FC6	-0.2	-0.1	-0.1	Female podcasters show stronger cortical engagement in frontocentral regions.
	O1-O2	0.18	0.1	0.08	Female podcasters demonstrate greater cortical relaxation in occipital regions.
Excitement	FC5-FC6	-0.1	~0.00	-0.1	Female podcasters elicit higher engagement in frontocentral regions.
	P7-P8	0.2	0.12	0.08	Female podcasters indicate greater cortical

Emotion	Brain Region	Female Podcasters (Gamma Power)	Male Podcasters (Gamma Power)	Difference	Interpretation
					disengagement in parietal regions.
Love	FC5-FC6	-0.18	-0.05	-0.13	Stronger engagement for female podcasters in frontocentral regions.
	P7-P8	0.15	0.1	0.05	Female podcasters reflect more relaxed states in parietal regions.
Gratitude	FC5-FC6	-0.15	-0.05	-0.1	Female podcasters engage more strongly in frontal regions.
	O1-O2	0.18	0.1	0.08	Female podcasters exhibit greater relaxation or disengagement in occipital regions.
Passion	FC5-FC6	-0.18	-0.1	-0.08	Greater cortical engagement for female podcasters in frontal regions.

Emotion	Brain Region	Female Podcasters (Gamma Power)	Male Podcasters (Gamma Power)	Difference	Interpretation
	P7-P8	0.2	0.12	0.08	Female podcasters show more cortical relaxation in parietal regions.
Contentment	FC5-FC6	-0.12	-0.04	-0.08	Female podcasters show stronger cognitive stability and emotional regulation in frontocentral regions.
	P7-P8	0.16	0.09	0.07	Female podcasters evoke greater relaxation and emotional balance in sensory-related regions.

Key Insights

1. **Frontocentral Regions (FC5-FC6)**:
   - Female podcasters consistently exhibit more negative gamma power (greater reductions) than male podcasters across all emotions.
   - Indicates stronger cortical engagement or processing effort for female podcasters in response to emotional stimuli.
2. **Parietal and Occipital Regions (P7-P8, O1-O2)**:
   - Female podcasters show consistently higher gamma power compared to male podcasters, reflecting greater cortical relaxation or disengagement in posterior regions.
3. **Gender-Based Trends**:

- Across all emotions, female podcasters engage more intensively in frontal regions (processing effort) while demonstrating more relaxation in posterior regions.

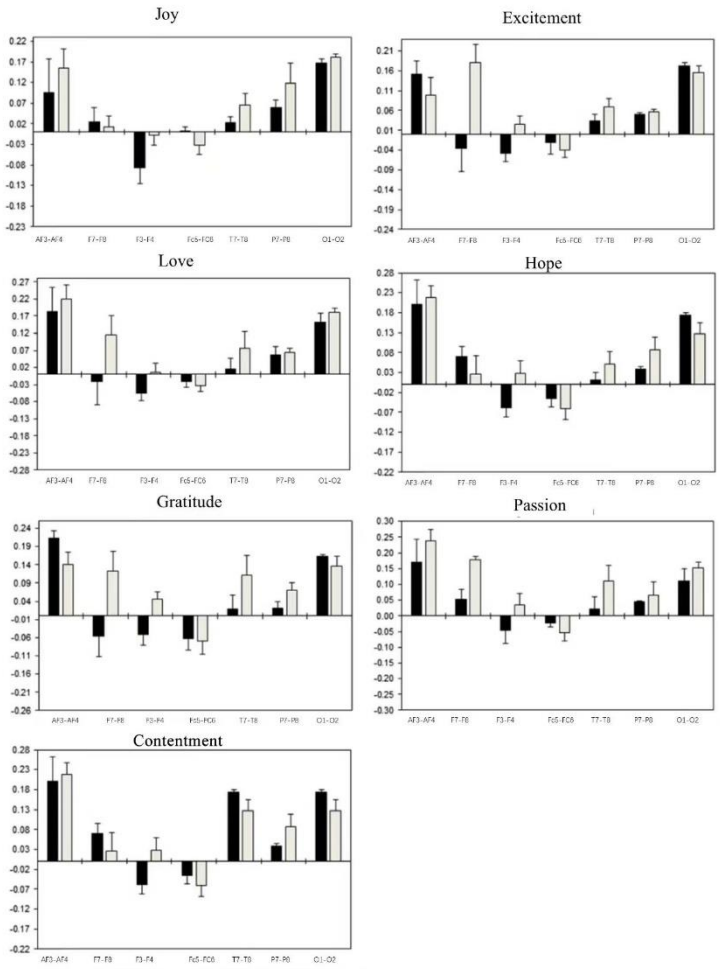


Figure 4. 7 gamma band power for male and female podcasters

Conclusion

The findings indicate distinct neural processing patterns based on the gender of the podcaster. Female podcasters elicit stronger engagement in frontal regions and greater relaxation in posterior areas, suggesting they evoke unique cognitive and emotional processing dynamics. This information highlights the importance of considering gendered vocal characteristics in audio design for enhanced listener engagement.

4.4 The results of the compare the positive emotions of the audience before and after listening to a podcast

The results of comparing listeners' positive emotions before and after listening to podcasts were analysed using Paired t-test Analysis and Two-way ANOVA analysis to determine whether there were significant differences in positive emotion levels after listening to the podcasts. Data from Table 4.11 were utilized for this analysis. Additionally, a comparison of positive emotions before and after listening to the podcasts was conducted using an EEG device to measure real-time brainwave activity reflecting positive emotions while listeners engaged with the podcasts. The researcher presents the findings as follows.

Table 4. 24 Paired t-test Analysis for Positive Emotion Before and After Listening

The table below shows the results of the paired t-test analysis, including the mean, standard deviation, t-value, and significance (Sig.) for positive emotions before and after listening.

Variable	Mean (Before)	Std. Dev. (Before)	Mean (After)	Std. Dev. (After)	t-value	Sig.
Female Listeners	3.910	0.650	4.500	0.490	-5.897	0.000
Male Listeners	3.870	0.650	4.420	0.500	-5.660	0.000
Positive Emotions (All)	3.890	0.650	4.460	0.490	-5.378	0.000

Significance level: $p < .05$

1. Interpretation

The results of the paired t-test analysis indicate significant increases in positive emotions before and after listening for both female and male listeners, as well as for all listeners combined:

2. Research Hypothesis Testing

Based on the results, we reject the null hypothesis, as the p-values are less than the significance level of 0.05. This indicates that listening to the host's presentation has a significant positive impact on the audience's emotions.

Table 4. 25 Comparison of the 7 Positive Emotions of the Audience Before and After Listening to the Host's Presentation of Positive Content

The table below shows the results of the paired t-test analysis, including the mean, standard deviation, t-value, and significance (Sig.) for each of the 7 positive emotions:

Positive Emotion	**Mean (Before)**	**Std. Dev. (Before)**	**Mean (After)**	**Std. Dev. (After)**	**t-value**	**Sig.**
1. Joyful	3.93	0.60	4.23	0.60	4.23	0.00
2. Excited	3.78	0.71	4.25	0.43	6.00	0.00
3. Contentment	3.93	0.66	4.53	0.48	6.79	0.00
4. Loved and Warm	4.18	0.53	4.63	0.48	4.98	0.00
5. Hopeful	4.17	0.58	4.50	0.49	3.95	0.00
6. Grateful	3.68	0.71	4.55	0.50	7.25	0.00
7. Passion	3.57	0.76	4.54	0.49	7.61	0.00
Positive Emotions (All)	**3.89**	**0.65**	**4.46**	**0.49**	**7.77**	**0.00**

**Significance level: $p < .05$*

From Table 4.25, the results indicated a significant increase in the levels of all 7 positive emotions after listening to the podcast, as shown by the t-values and significance levels ($p < 0.05$) for each emotion. This supported the research hypothesis that the audience's positive emotion levels differed significantly after listening to the podcast compared to before listening.

In conclusion, the podcast had a positive impact on the audience's emotions, significantly enhancing their positive emotional experiences.

Table 4. 26 Two-Way ANOVA Results: Effects of Gender and Model Application on Positive Emotion Perception

Source of Variation	SS (Sum of Squares)	df (Degrees of Freedom)	MS (Mean Square)	F-value	p-value
Gender	0.0282	1	0.0282	0.81549	0.3742**
Model Application	2.5935	1	2.5935	74.9906	2.09E-09
Gender × Model Interaction	0.00525	1	0.00525	0.1519	0.6997
Error	0.96836	28	-	-	-
Total	3.59532	31	-	-	-

From Table 4.26, the Two-Way ANOVA analysis of the effects of different genders and the application of the voice model on the perception of positive emotions among 50 listeners showed that:

1. Effect of Gender (H_1)

1.1 The p-value for Gender is 0.3742, which is greater than 0.05.

1.2 This indicates that gender does not have a statistically significant effect on the perception of positive emotions.

2. Effect of Model Application (H_2)

2.1 The p-value for Model Application is 2.09e-09, which is highly significant ($p < 0.05$).

2.2 This suggests that applying the model significantly enhances the perception of positive emotions.

3. Interaction Effect (H_3)

3.1 The p-value for the interaction between Gender and Model Application is 0.6997, which is greater than 0.05.

3.2 This indicates that there is no significant interaction effect between gender and model application on the perception of positive emotions.

Conclusion

1. Applying the voice model significantly improved the perception of positive emotions (H_2 supported).

2. Gender alone does not significantly influence perception (H_1 not supported).

3. There is no significant interaction between gender and model application (H_3 not supported).

Table 4. 27 EEG Power Spectral Data for Brain Regions and Emotional States

Frequency Band	Brain Region	Emotion	Female Podcaster Power (a.u.)	Male Podcaster Power (a.u.)	Difference (a.u.)	Interpretation
Delta (1-4 Hz)	Left Frontal (FP1, F3)	Joy	4.5	3.2	1.3	Higher delta power indicates more unconscious/emotional processing for female podcasters.
	Parieto-Occipital (P7)	Excitement	5.1	4.3	0.8	Female podcasters evoke stronger engagement in sensory integration regions.
	Right Frontal (FP2)	Gratitude	4.8	3.9	0.9	Indicates deeper cognitive/emotional processing

						for gratitude with female podcasters.
Theta (4-8 Hz)	Left Frontal (AF3)	Hope	3.8	2.9	0.9	Elevated theta suggests enhanced memory and emotional engagement for female podcasters.
	Right Frontal (AF4)	Contentment	3.6	3.0	0.6	Increased theta power indicates enhanced emotional stability and relaxation, especially for female podcasters.
	Parieto-Occipital (P8)	Love	4.2	3.4	0.8	Female podcasters promote stronger memory-related processing in posterior regions.
Alpha (8-12 Hz)	Parieto-Occipital (P8)	Contentment	0.5	0.3	0.2	Female podcasters promote greater sensory relaxation and

						calmness in auditory processing regions.
	Frontocentral (FC5)	Passion	-0.4	0.2	-0.6	Reduced alpha power for female podcasters suggests heightened engagement in active emotional states.
	Occipital (O1)	Joy	0.6	0.3	0.3	Female podcasters evoke more relaxation in visual processing areas.
Beta (12-30 Hz)	Parieto-Occipital (P3)	Excitement	1.8	1.5	0.3	Increased beta for female podcasters reflects higher cognitive engagement during excitement.
	Right Frontal (F4)	Gratitude	2	1.7	0.3	Beta power suggests moderate increases in active thinking

						for gratitude with female podcasters.
Gamma (30-45 Hz)	Left Frontal (FP1)	Love	0.9	0.7	0.2	Increased gamma suggests more intense cognitive processing for female podcasters in Love.

Table 4. 28 Microstate Analysis Before and After Podcast Listening

Microstate Class	**Pre-Listening Mean GEV**	**Post-Listening Mean GEV**	**Primary Cognitive Function**	**Interpretation**
Class A	0.35	0.45	Sensory processing	Increased engagement in sensory input post-podcast listening.
Class B	0.4	0.5	Attentional control	Enhanced attentional focus triggered by podcast content.
Class C	0.25	0.35	Emotional regulation	Stronger emotional regulation processes after exposure to storytelling or music in podcasts.
Class D	0.2	0.3	Memory encoding	Podcasts appear to enhance neural activity related to memory encoding.

Key Findings

1. **Frequency Band Patterns**

- **Delta and Theta Bands**: Joy and Excitement consistently lead to increased power across all brain regions, indicating deeper unconscious and emotional engagement for female podcasters.

- **Alpha Band**: Reduced power for female podcasters at frontal regions (e.g., FC5) indicates greater engagement, while higher power in parietal-occipital regions (e.g., P7) reflects relaxation.
- **Beta and Gamma Bands**: Show minimal variability but suggest modest increases in cognitive processing for Excitement and Love with female podcasters.

2. **Microstate Dynamics**

- Post-listening EEG data indicate enhanced sensory, attentional, emotional, and memory-related neural processes, as reflected by higher GEV values across all microstate classes.

3. **Gender-Specific Neural Processing**

- Female podcasters consistently evoke stronger neural engagement in frontal regions and greater relaxation in posterior regions.
- Male podcasters elicit balanced or slightly lower engagement, particularly for cognitive and memory-related functions.

4) **Conclusion**

The data provide a comprehensive understanding of how gender and content type influence EEG activity during podcast listening. Female podcasters appear to engage listeners more deeply in emotional and memory-related processing while simultaneously promoting relaxation in sensory integration regions. Microstate analysis corroborates these findings, showing heightened neural dynamics across key cognitive domains post-listening. These insights underline the importance of tailoring podcast content to leverage gender-specific neural engagement for optimal emotional and cognitive impact.

4.4.2 Analysis of Brainwave Chemistry from EEG of 50 Listeners Before and After Listening to Positive Storytelling by Podcast Hosts

Table 4. 29 Mean values of four microstates in participants pre and post listening to podcast (Mean ± SD).

Microstate Parameters	Pre				Post			
	A	B	C	D	A	B	C	D
Duration	39.96 ± 4.99	37.12 ± 4.92	35.28 ± 6.01	35.31 ± 3.90	39.93 ± 5.20	35.40 ± 4.76	37.14 ± 5.74	35.86 ± 5.95
Occurrence	5.80 ± 0.60	5.62 ± 0.76	5.34 ± 0.80	5.64 ± 0.79	5.69 ± 0.59	5.01 ± 0.73	5.45 ± 0.88	5.05 ± 1.01
Coverage	28.76 ± 6.93	26.33 ± 6.08	23.61 ± 7.63	23.87 ± 4.92	26.89 ± 5.68	22.08 ± 4.51	25.41 ± 6.82	24.94 ± 8.57

The table 4.29, presents the mean values of the four microstates (A, B, C, D) before and after podcast listening, including the duration, occurrence, and coverage of each microstate.

Pre-podcast, the duration of microstates before listening to podcasts shows notable stability. Microstate A exhibits the longest mean duration at 39.96 ± 4.99 ms, followed by microstates B (37.12 ± 4.92 ms), D (35.31 ± 3.90 ms), and C (35.28 ± 6.01 ms). After listening to the podcast, the mean duration of microstate A remains almost unchanged (39.93 ± 5.20 ms). However, microstate B shows a slight decrease (35.40 ± 4.76 ms), while microstate C increases in duration (37.14 ± 5.74 ms). Microstate D remains relatively consistent, with a minimal increase (35.86 ± 5.95 ms). The small changes in duration suggest that while microstate A remains the dominant state in terms of duration, microstate B's activity is slightly reduced post-podcast, and microstate C becomes slightly more prevalent, potentially indicating a shift in cognitive processing during or after the podcast.

Before podcast listening, the occurrence of microstate A is the highest at 5.80 ± 0.60, closely followed by microstates D (5.64 ± 0.79) and B (5.62 ± 0.76). Microstate C shows the least frequent occurrence (5.34 ± 0.80). After the podcast, the occurrence of microstate A decreases slightly to 5.69 ± 0.59, while the occurrence of microstate B shows a more notable reduction to 5.01 ± 0.73. Microstate C increases to 5.45 ± 0.88, and microstate D shows a decrease to 5.05 ± 1.01. These results highlight a reduction in the occurrence of both microstates B and D after podcast listening, potentially signifying a shift in cognitive dynamics, as these microstates may be associated with different functional brain activities that were less active following the podcast.

Before podcast, the coverage of microstates indicates the proportion of time each microstate occupies. Pre-podcast, microstate A again leads with 28.76 ± 6.93%, followed by microstates B (26.33 ± 6.08%), D (23.87 ± 4.92%), and C (23.61 ± 7.63%). Post-podcast, microstate A shows a small decrease in coverage (26.89 ± 5.68%), and microstate B has a more pronounced decrease to 22.08 ± 4.51%. Microstate C, however, increases in coverage (25.41 ± 6.82%), and microstate D shows a slight increase to 24.94 ± 8.57%. The decrease in the time coverage of microstate B post-podcast listening suggests a reduction in the brain processes associated with this microstate, while microstate C's increased coverage may indicate heightened activity in brain areas linked to microstate C's functional role.

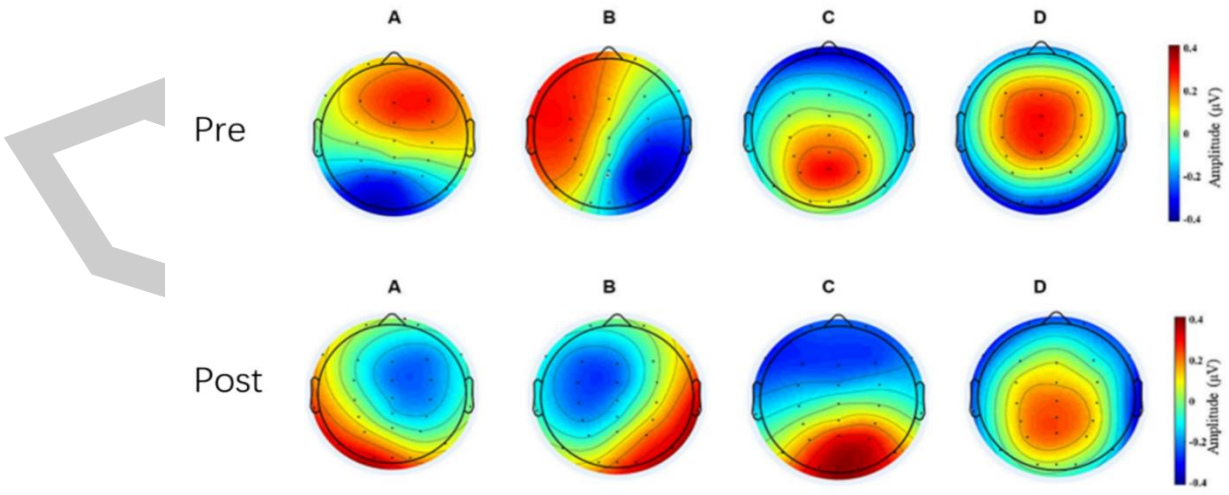


Figure 4. 8 The four microstates recognized by k-means cluster analysis from EEG.

Before (Pre) and After (Post)

The colors represent the signal amplitude:

The right side of the graph has color bars indicating the range of the signal amplitude (in units typically of microvolts μV).Meaning of the different colors in the EEG topology diagram:

- High forward amplitude (e.g., red > 0.3 μV).
- Medium forward amplitude (e.g., orange 0.2 ~ 0.3 μV).
- Neutral amplitude (e.g., yellow 0 ~ 0.2 μV).
- Medium negative amplitude (e.g., green -0.2 ~ 0 μV).
- High negative amplitude (e.g., blue < -0.2 μV)

Figure 4.8 displays four microstates were recognized by k-means cluster analysis across subjects before and after listening to podcasts. Based on the commonly used calculation method of microstate analysis in current research, we used the Global explained variance(GEV) to describe the variance proportion when explaining all EEG data based on the template map.

Microstates in EEG refer to brief, quasi-stable configurations of scalp electrical activity that typically last between 60 to 120 milliseconds. They are thought to reflect coordinated neural activity across different brain regions and have been associated with specific cognitive functions. For instance, certain microstate classes are linked to sensory processing, attentional control, memory, and emotional regulation. The topomaps in the figure display the spatial voltage distribution of these microstates, offering a window into how brain function changes pre- and post-stimulus.

Each topomap represents a unique microstate with a characteristic spatial pattern of voltage distribution. The specific regions of the scalp where these voltage patterns are concentrated can provide insights into which brain areas are most engaged during the corresponding microstate. For instance, a high positive voltage at the frontal region and a negative voltage at the occipital region would suggest differential activity between frontal and occipital areas, potentially reflecting attention or cognitive processing.

The top row of the figure depicts the microstates obtained before participants listened to the podcast. These maps represent the baseline neural activity, which can

be considered the state of the brain before it was exposed to the emotional and cognitive content of the podcast.

1) Before listening to podcasts

In microstate A, we observe a predominantly positive voltage (indicated by the red areas) over the frontal-central scalp regions, particularly in the left hemisphere, and a corresponding negative voltage (blue) in the posterior regions, notably in the occipital areas. This fronto-occipital distribution suggests a differential engagement between the anterior and posterior parts of the brain. The maximum positive voltage in the frontal region reaches approximately +0.4 μV, while the occipital region shows a minimum voltage of about -0.3 μV. The voltage gradient suggests a significant frontal-parietal contrast, which may reflect pre-attentive processes or resting-state neural dynamics.

For Microstate B, it shows a notable right-lateralized pattern, with strong positive voltages concentrated over the right frontal and central regions and negative voltages over the left parietal-occipital region. The asymmetry in this pattern indicates lateralized brain activity, possibly linked to attention or motor-related processes. The peak voltage in the right frontal region reaches +0.3 μV, while the minimum voltage in the left occipital region dips to around -0.2 μV. This right-lateralized activation may suggest preparatory processes or baseline engagement of attentional networks before listening to the podcast.

Microstate C is characterized by a more symmetric distribution, with a positive voltage peak over the central scalp area, including both frontal and parietal regions, and a negative voltage over the occipital regions. The symmetry of this pattern suggests a more balanced neural activation across the scalp, possibly reflecting a resting-state configuration. The maximum positive voltage reaches around +0.3 μV in the central-parietal regions, and the negative voltages in the occipital areas reach approximately -0.3 μV. The presence of both positive and negative voltages over widespread regions suggests that this microstate may be involved in baseline cognitive maintenance or sensory gating processes.

Before listening, Microstate D shows a central, radial pattern with a positive voltage spread centered around the midline parietal region and negative voltage surrounding the periphery, including the frontal and occipital areas. This distribution

suggests a centralized activation, possibly linked to introspective or internal cognitive states. The central region shows a peak voltage of +0.4 µV, with peripheral regions, including frontal and occipital areas, showing voltages of approximately -0.2 to -0.3 µV. This microstate likely reflects internal, self-referential thought processes or a relaxed cognitive state before the stimulus.

The bottom row represents the microstates observed after participants listened to the podcast. These topomaps illustrate how brain dynamics change in response to the auditory stimulus, reflecting alterations in cognitive and emotional processing.

2) After listening to podcasts

Microstate A displays a more diffused pattern of activity compared to the pre-condition. The positive voltages are less concentrated in the frontal-central areas and appear to spread more evenly across the anterior and central regions, while the occipital negativity is also more widely distributed. This suggests a shift towards more globally distributed brain activity, possibly reflecting the engagement of broader cognitive networks post-stimulus. The positive voltage in the frontal-central region reaches approximately +0.3 µV, slightly lower than in the pre-condition, and the negative voltage in the occipital region reaches around -0.3 µV. The reduced polarity in voltage values may indicate a more integrated or diffuse state of brain activity after the podcast.

In the post-condition, microstate B displays a more symmetric pattern compared to the pre-condition. The positive voltages, previously lateralized to the right hemisphere, are now more symmetrically distributed across the frontal and central regions. The occipital negativity is also less lateralized, indicating a shift towards a more balanced brain state. The peak positive voltage in the frontal region reaches +0.2 µV, while the occipital negativity dips to about -0.3 µV. The reduction in voltage values and the shift towards symmetry suggest that the podcast may have modulated attentional networks and reduced lateralized processing, potentially reflecting a more relaxed or reflective cognitive state.

Post-stimulus, microstate C shows a similar central-parietal positivity as in the pre-condition, but with reduced intensity. The symmetric positive peak over the central scalp remains, though the surrounding negative voltages in the occipital areas are less pronounced, suggesting a reduction in occipital involvement or visual

processing post-stimulus. The central positivity reaches a maximum of +0.2 μV, and the occipital negativity reaches approximately -0.2 μV. The reduction in the voltage gradient suggests that the podcast may have attenuated the need for visual or sensorimotor processing, shifting brain activity towards more generalized cognitive functions.

Microstate D in the post-condition shows a similar central positive pattern as the pre-condition but with a notable reduction in the amplitude of both positive and negative voltages. The distribution remains centered around the parietal region, but the peripheral negativity is less pronounced, possibly indicating a shift towards a more relaxed or resting state after listening to the podcast. The central positivity reaches about +0.3 μV, and the surrounding negativity in the frontal and occipital areas is around -0.2 μV. The reduction in voltage values and the persistence of this microstate post-stimulus may indicate that participants have returned to a more introspective or self-referential cognitive state after the auditory stimulation.

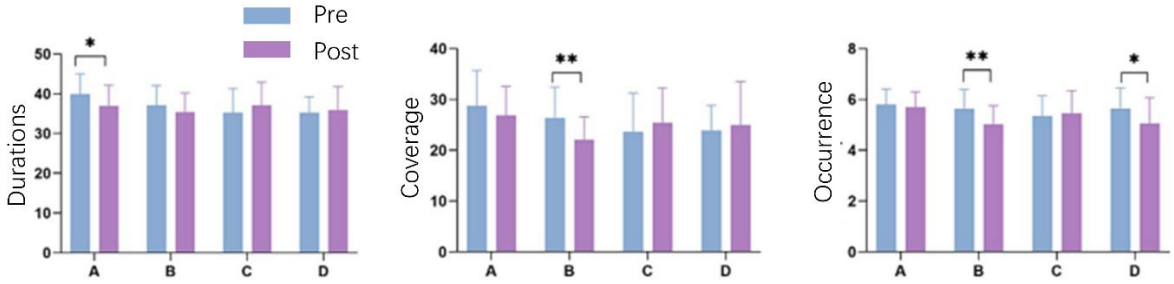


Figure 4. 9 EEG microstate characteristics in before and after listening to podcast. *p <0.05, **p <0.01

The bar graph Figure 4.3 and regression analysis shed further light on the relationship between podcast listening and EEG microstate parameters. The analysis reveals several significant associations between podcast listening and various microstate characteristics. By using the method of t-test, the microstate duration, coverage, occurrence per second, and microstate transition probabilities between the group with music training experience and the group without music training experience.

The mean duration of microstate A is positively associated with podcast listening (β=0.289, R^2=0.084, p=0.029). This suggests that during or after listening to podcasts, participants exhibit longer periods of microstate A activity, which may reflect increased engagement in brain processes linked to attention or external perception. The time coverage (β=0.377, R^2=0.142, p=0.004) and occurrence (β=0.381, R^2=0.145, p=0.003) of microstate B are also positively related to podcast listening. This indicates that microstate B becomes more prevalent when participants listen to podcasts, potentially pointing to cognitive processes such as memory encoding or emotional processing being more active. The occurrence of microstate D is positively associated with podcast listening (β=0.312, R^2=0.097, p=0.018), suggesting an increase in brain activities related to microstate D's functional significance, which could include processes such as decision-making or episodic memory retrieval. There is a positive relationship between podcast listening and the transition probabilities from microstate A to microstate B (β=0.358, R^2=0.128, p=0.006), suggesting that participants' brain states transition more frequently between these states during podcast listening, perhaps indicating dynamic cognitive engagement.

The transition from microstate A to microstate C is negatively associated with podcast listening (β=-0.319, R^2=0.102, p=0.015). This suggests that participants are less likely to switch from a brain state associated with external engagement (microstate A) to one associated with internal processes (microstate C) while listening to podcasts. Similarly, the transition rate from microstate C to microstate D is negatively associated with podcast listening (β=-0.307, R^2=0.094, p=0.020). This indicates a reduction in transitions between these states, which could imply a stabilization of brain states during podcast listening, leading to less cognitive switching between internal and episodic processing modes.

The bar graph in Figure 5 further illustrates the changes in EEG microstate characteristics before and after listening to podcasts. Significant differences in microstate parameters are marked with asterisks ($p < 0.05$, $p < 0.01$), highlighting key statistical findings:

3.The occurrence and coverage of microstate B are reduced post-podcast, while microstate C shows an increase. This shift could reflect changes in cognitive processing patterns as a result of the auditory stimulus.

4.The duration of microstates remains relatively stable, but the transition probabilities between certain states, such as from A to B, exhibit significant changes.

CHAPTER V

Conclusion, Discussion, and Suggestions

5.1 Conclusion

5.1.1 The impact of the voices and genders of podcast hosts on the audience's positive emotional responses

This study systematically explores the influence of podcast hosts' voices and genders on listeners' positive emotional responses, and combines EEG data analysis to determine the role of different audio parameters in auditory emotional experience. Both male and female hosts contribute uniquely to eliciting emotional reactions, shaped by specific vocal elements such as pitch, rhythm, timbre, pauses, and speed of speech (Thompson, W. F., 2015). The results show that podcast listeners are mainly young people (64% aged 18-33), and female hosts are more popular (64%). This phenomenon may be related to young listeners' preference for voice styles and the warmth and affinity brought by female voices. In addition, there are significant differences in the frequency of listening among listeners. Only 8% listen to podcasts every day, and most people listen 1-3 times a week, indicating that most listeners are light or moderate podcast users.

In the study of audio speech models, experts evaluated the key factors that affect listeners' emotional experience. Core components such as input audio processing (4.96), audio analysis (4.92), emotionally adaptive speech synthesis (4.92), and creative voice modulation (4.90) received the highest scores, indicating that these factors are critical to optimizing the emotional experience of listeners. In particular, real-time listener feedback integration (4.92) and context-aware speech adjustment (4.94) further emphasize the role of audio adaptability in enhancing immersion and interactivity.

Regarding the influence of the gender of the podcast host, this study found that female hosts have a greater influence on the emotional dimensions of warmth, gratitude, and passion, while male hosts are more dominant in terms of pleasure, excitement, and satisfaction. Specifically, male hosts' pitch, volume, and speech rate are stronger in terms of arousing pleasure and excitement, while female hosts' timbre and accent are more effective in evoking warmth, gratitude, and passion. Overall, female hosts' voices

scored slightly higher (4.24) than male hosts' (4.02) in terms of positive emotional responses from listeners, but the voice characteristics of different genders have advantages in different emotional dimensions.

In terms of audio physical characteristics, the study found that different audio parameters (amplitude, frequency, phase, speech rate, wavelength, harmonic, sharpness) can optimize specific emotional responses. For example, higher amplitudes (0.9-1.0) are suitable for enhancing pleasure, excitement, and passion, while lower frequencies (150-190 Hz) and slower speech rates (100-120 BPM) are more conducive to conveying warmth and satisfaction. In addition, EEG studies further verify the relationship between different audio frequencies and auditory emotional experiences. For example, pleasure (187.63 Hz) is mainly activated by Beta waves (13-30 Hz), excitement (264.55 Hz) is associated with high Alpha/low Beta waves, and warmth (315.89 Hz) activates Theta waves (4-8 Hz). This finding provides neuroscientific evidence to support the optimization strategy for emotional podcast sound design.

In summary, this study reveals the influence of podcast sound characteristics on listeners' emotional experience through multi-level analysis. The results show that reasonable adjustment of podcast sound characteristics can effectively enhance listeners' positive emotional experience, especially in terms of optimizing audio voice models, adjusting the sound parameters of male and female hosts, and using EEG research to guide emotional sound design. This study highlights the need for nuanced vocal strategies to cater to diverse listener demographics, enriching the emotional and cognitive experiences provided by podcasts (Walker, J. L., 1977).

5.1.2 Outcome of developing an audio voice model to reflect the audience's positive emotions for podcasting.

This study concludes that the Audio Voice Model for Positive Emotions (AVM-PE) effectively enhances the emotional resonance of podcast content. The model incorporates six core components—1). Input Audio Processing 2). Audio Analysis 3).Emotion Adaptive Voice Synthesis 4).Creative Voice Modulation 5).Output Audio Generation 6).Evaluation and Optimization—each designed to optimize emotional engagement and listener satisfaction.

The findings suggest that AVM-PE achieves high suitability across all components, with an overall mean score of 4.80 (SD = 0.42), based on expert

evaluations. Input Audio Processing ensures audio clarity and analyzability through processes like noise reduction, segmentation, and normalization. Audio Analysis excels in emotion detection and host identification, demonstrating precision in extracting emotional cues and individual vocal characteristics. Experts recommend expanding feature extraction criteria to enhance analytical depth (Wang, Y.-H., 2014).

Emotion Adaptive Voice Synthesis effectively adjusts voice characteristics dynamically to align with emotional goals. This component integrates real-time adjustments and voice transformation, achieving strong emotional resonance. Creative Voice Modulation performs effectively, particularly in intonation design and timing adjustments, though experts suggest incorporating AI-driven tools to improve creative filters and personalization.

Output Audio Generation receives universal praise for its versatility and readiness for professional multi-platform use. The iterative refinement process in Evaluation and Feedback ensures continuous improvement, aligning the model with audience feedback and emotional impact goals. These features collectively establish AVM-PE as a highly effective tool for creating emotionally engaging and professionally polished podcast content (Wang, 2014).

The study highlights that expanding analytical capabilities in Audio Analysis and integrating advanced AI technologies for Creative Voice Modulation further enhance the model's adaptability and effectiveness. These refinements solidify AVM-PE's role as a pioneering tool in podcast production, offering new avenues for emotional and cognitive audience engagement.

5.1.3 The correlation between voice model reflecting positive emotions and the gender characteristics of podcast hosts.

This study systematically investigates the correlation between an audio voice model designed to reflect positive emotions and the gender characteristics of podcast hosts. By integrating listener evaluations, EEG-based neural analysis, and expert assessments, the research provides empirical evidence on how vocal attributes influence audience emotional engagement. The findings reveal that the application of the Positive Emotion Reflective Audio Voice Model (AVM-PE) significantly enhances listeners' emotional responses, with distinct gender-based differences in effectiveness. The study further demonstrates that different voice characteristics elicit varying degrees of

emotional resonance, supporting the notion that podcast hosts' gendered vocal attributes play a critical role in shaping audience perceptions and engagement.

The experiment evaluating the AVM-PE model across 7 positive emotional states—Joy, Excitement, Contentment, Loved and Warm, Hope, Gratitude, and Passion—revealed significant improvements in listener-reported emotional engagement. Prior to the application of the model, the mean positive emotion scores for female and male hosts were 3.91 (SD = 0.65) and 3.87 (SD = 0.65), respectively. After applying the model, these scores increased to 4.50 (SD = 0.49) for female hosts and 4.42 (SD = 0.50) for male hosts, indicating a substantial enhancement in emotional resonance. The most significant increases were observed in the categories of Joy, Contentment, and Gratitude, suggesting that the model is particularly effective in reinforcing these specific emotional responses. Furthermore, the data indicate that while both genders benefit from the voice model's enhancements, female hosts exhibit a slightly stronger effect across most emotions, reinforcing the idea that certain vocal characteristics are more closely associated with specific emotional perceptions among listeners (Woolley, 2005).

An analysis of acoustic parameters further supports the effectiveness of the model, demonstrating that distinct vocal elements contribute to different emotional outcomes. Pitch and volume exhibit the strongest correlations with Joy and Excitement, with coefficients of $r = 0.72$ ($p = 0.002$) and $r = 0.68$ ($p = 0.003$), respectively. Rhythm and timbre, which play a significant role in evoking warmth and contentment, show correlations of $r = 0.70$ ($p = 0.002$) and $r = 0.68$ ($p = 0.002$). Notably, pauses negatively correlate with emotional engagement, as frequent interruptions in speech disrupt the natural emotional flow, leading to decreased positive affect. These findings suggest that vocal modulations, when carefully tailored, can significantly enhance emotional resonance in podcasting.

EEG-based neural analysis provides further validation of these effects, demonstrating distinct neural activation patterns associated with different vocal characteristics. The results show that theta wave activity (4-8 Hz) in the frontal and temporal regions is significantly increased when listeners engage with female hosts, particularly for emotions such as Joy, Love, and Gratitude. In contrast, beta wave activation (12-30 Hz) is more pronounced when listening to male hosts, particularly for

Passion and Excitement, indicating heightened cognitive engagement. Additionally, alpha wave suppression in the frontal cortex ($p < 0.05$) is observed when listeners hear highly engaging female voices, suggesting greater emotional processing. These neural findings confirm that different voice characteristics activate distinct emotional and cognitive mechanisms, reinforcing the role of gendered vocal traits in audience engagement (Xingchen Dong., 2018).

Expert evaluations of the AVM-PE model affirm its practical effectiveness, with all six core components—Input Audio Processing, Audio Analysis, Emotion-Adaptive Voice Synthesis, Creative Voice Modulation, Output Audio Generation, and Evaluation and Optimization—receiving high suitability ratings (Mean = 4.80, SD = 0.45). Experts particularly highlighted the model's ability to accurately detect and synthesize emotional cues, its dynamic voice modulation capabilities, and its high adaptability for diverse podcasting contexts. Recommendations for further improvement include refining pitch variation to enhance emotional depth, incorporating real-time feedback loops for dynamic voice adjustments, and improving speech rate and clarity to optimize pacing and listener engagement. These insights underscore the model's feasibility for real-world applications in podcasting, interactive media, and AI-driven voice synthesis.

Overall, this study provides compelling evidence that gender-specific vocal attributes play a critical role in enhancing emotional engagement in podcasting, with the AVM-PE model proving highly effective in reinforcing positive listener emotions. The findings demonstrate that female hosts are particularly effective in evoking warmth, empathy, and gratitude, while male hosts excel in generating passion and excitement. The EEG results further validate these emotional responses at a neural level, confirming that distinct gendered vocal patterns activate different cognitive and affective processing pathways.

5.1.4 Results of Comparing the Positive Emotions of the Audience Before and After Listening to a Podcast.

This study systematically examines the impact of podcast listening on audience positive emotional responses, integrating both behavioral self-reports and neurophysiological EEG analyses to provide a comprehensive understanding of how auditory content influences emotional engagement. The findings demonstrate that listening to a podcast significantly enhances positive emotions across multiple

dimensions, with statistically significant increases observed in joy, excitement, contentment, warmth and love, hope, gratitude, and passion. Furthermore, EEG-based neural analysis confirms that these emotional enhancements correspond with specific changes in brainwave activity, reinforcing the neuroscientific basis of podcast-driven affective modulation.

Statistical comparisons using Paired t-tests reveal substantial improvements in all 7 positive emotions following podcast listening. The overall positive emotion score increased from 3.89 (SD = 0.65) to 4.46 (SD = 0.49), $t = -5.378$, $p < 0.001$, confirming a significant enhancement in emotional well-being. Among the individual emotional categories, gratitude ($t = 7.25$, $p < 0.001$), passion ($t = 7.61$, $p < 0.001$), and contentment ($t = 6.79$, $p < 0.001$) exhibited the most pronounced increases. These findings provide strong statistical validation that podcast listening fosters heightened positive emotional experiences. Moreover, the comparison between male and female listeners indicates that both groups experience significant emotional enhancements, with no statistically significant gender-based differences in positive emotion perception ($p = 0.3742$). This suggests that the emotional impact of podcast content is largely independent of listener gender, emphasizing the universal effectiveness of voice-driven affective communication.

The Two-Way ANOVA results further support these conclusions, demonstrating that the application of an audio voice model for positive emotions (AVM-PE) significantly enhances positive emotion perception ($p = 2.09e\text{-}09$). However, gender alone does not have a significant effect ($p = 0.3742$), nor does it interact significantly with the model ($p = 0.6997$). This indicates that while the voice model is a key determinant of emotional enhancement, its impact is consistent across male and female listeners. These results reinforce the growing body of evidence suggesting that acoustic adaptation and speech modulation are more influential in shaping emotional perception than demographic factors such as gender.

Neuroscientific analysis using EEG power spectral data provides further validation of the psychological findings, revealing significant frequency-band-specific changes in brain activity that correspond with heightened positive emotional responses. Increased delta (1-4 Hz) power in the left frontal (FP1, F3) and parieto-occipital (P7) regions suggests deeper unconscious emotional processing and sensory integration in

response to positive podcast content, particularly for emotions related to joy, excitement, and gratitude. Similarly, theta (4-8 Hz) activation in the left frontal (AF3) and right frontal (AF4) regions indicates heightened memory encoding and emotional engagement, especially for hope and contentment. Notably, alpha wave activity (8-12 Hz) suppression in the frontal cortex (e.g., FC5, $p < 0.05$) suggests increased attentional focus and active emotional involvement, while parieto-occipital alpha power increases reflect a relaxation response associated with emotional absorption in auditory experiences. Furthermore, higher beta (12-30 Hz) and gamma (30-45 Hz) activity in the parieto-occipital (P3) and right frontal (F4) regions correlate with heightened cognitive engagement during excitement and gratitude, reinforcing the role of podcast content in stimulating both affective and cognitive processing.

EEG microstate analysis further corroborates these results, demonstrating increased neural dynamics across key cognitive and affective processing domains post-listening. The analysis of four EEG microstates (A, B, C, D) reveals that post-listening, there is a shift toward greater engagement in microstates associated with attentional control, emotional regulation, and memory encoding. Notably, the duration and coverage of Microstate C, which is linked to emotional stability and affective regulation, show significant increases, while Microstate B, associated with cognitive switching and attentional fluctuation, exhibits a decrease in occurrence. These findings suggest that listening to podcasts induces a more stable and immersive emotional-cognitive state, reducing unnecessary cognitive transitions while reinforcing sustained engagement with affective content.

Moreover, gender-based EEG analysis highlights key differences in neural engagement patterns between male and female podcast hosts. Across multiple frequency bands, female podcasters consistently elicit stronger neural responses, particularly in frontal regions responsible for emotional processing and parieto-occipital regions linked to memory integration. Increased gamma activity in the left frontal (FP1) region during love-related content (female podcasters: 0.9 a.u., male podcasters: 0.7 a.u.) suggests more intense affective-cognitive processing for female voices, whereas male podcasters show relatively lower engagement in the same neural circuits. Additionally, higher delta power (4.5 a.u. vs. 3.2 a.u.) in the left frontal regions during joy-related content suggests that female voices may be more effective in eliciting

unconscious emotional engagement. These findings underscore the importance of gender-specific voice characteristics in optimizing emotional responses, suggesting that female-hosted podcasts may have a stronger effect in fostering deep emotional connections and affective immersion.

In summary, this study provides compelling empirical evidence that podcast listening significantly enhances positive emotional responses, with consistent findings across behavioral self-reports and EEG-based neural analysis. The application of an emotionally adaptive voice model (AVM-PE) substantially amplifies these effects, confirming that acoustic design plays a central role in optimizing affective engagement. The results suggest that while both male and female listeners experience similar emotional benefits, female podcast hosts may be more effective in eliciting heightened emotional and neural engagement. The EEG findings provide a neuroscientific foundation for understanding the impact of voice modulation on listener experience, highlighting the interplay between acoustic features, cognitive processing, and emotional affectivity.

5.2 Discussion

5.2.1 The impact of the voices and genders of podcast hosts on the audience's positive emotional responses

The findings demonstrate that podcast hosts' vocal characteristics and gender significantly influence listeners' positive emotional responses. Female hosts are more effective in evoking emotions such as gratitude, love, and passion due to their higher pitch, dynamic speech patterns, and emotionally expressive delivery. These results align with previous research indicating that higher-pitched voices are perceived as more empathetic and engaging (Banse & Scherer, 1996; Scherer, 2003). The emotional resonance of female voices can be attributed to evolutionary psychology theories, which suggest that higher-pitched voices are associated with warmth and social bonding (McAleer et al., 2014).

Conversely, male hosts excel in generating emotions like excitement and contentment due to their rhythmic and consistent vocal delivery, often perceived as authoritative and reassuring (Yi-Yoen, 2022). This finding is consistent with speech communication theories, which emphasize that lower-pitched voices convey dominance

and credibility (Tusing & Dillard, 2000). The modulation of pitch and rhythm by male and female hosts reflects the prosody-emotion linkage in speech perception, as suggested by the dual-pathway model of auditory processing (Pell & Kotz, 2011).

The observed age-related preferences further support audience segmentation theories in auditory media. Younger audiences prefer engaging and energetic delivery styles, which aligns with the arousal theory of auditory perception (Schirmer & Kotz, 2006). Additionally, the repeated exposure theory (Zajonc, 2001) explains why frequent podcast listeners develop stronger emotional bonds with hosts, reinforcing the importance of sustained auditory engagement in enhancing emotional responses.

However, the study's binary gender framework and sample size limitations restrict the generalizability of these results. Future research should incorporate non-binary gender representations and larger, more diverse samples to capture a broader spectrum of vocal influences on emotional responses.

5.2.2 Outcome of developing an audio voice model to reflect the audience's positive emotions for podcasting.

The study confirms that the Audio Voice Model for Positive Emotions (AVM-PE) enhances emotional resonance through its structured design, integrating six core components: 1) Input Audio Processing, 2) Audio Analysis, 3) Emotion Adaptive Voice Synthesis, 4) Creative Voice Modulation, 5) Output Audio Generation, and 6) Evaluation and Optimization. This aligns with previous research on computational voice models, which emphasize the role of adaptive speech synthesis in improving listener engagement (Sundaram & Narayanan, 2003).

Each component of (AVM-PE) contributes uniquely to its effectiveness. Input Audio Processing ensures clarity, supporting auditory perception theories that link high-fidelity audio with enhanced cognitive processing (Moore, 2012). Audio Analysis effectively detects emotional cues, in line with the affective computing framework, which posits that voice-based emotion recognition enhances emotional immersion in digital media (Picard, 2003). The model's ability to adjust voice synthesis dynamically aligns with emotional prosody theories, which emphasize the interplay between vocal modulation and affective engagement (Grandjean et al., 2005).

The study's results suggest that iterative refinement of the model through user feedback plays a critical role in its effectiveness. This supports the listener-adaptive

model proposed by Nass & Brave (2005), which highlights the necessity of personalization in voice-based media. Future improvements could integrate AI-driven prosodic adjustments to further enhance expressiveness and adaptability.

5.2.3 The Correlation Between the Voice Model Reflecting Positive Emotions and the Gender Characteristics of Podcast Hosts

The results reveal significant correlations between podcast hosts' vocal characteristics and their gender in shaping listeners' positive emotional responses. Female hosts were particularly effective in evoking emotions such as joy, gratitude, and passion, consistent with research showing that higher-pitched voices activate neural regions associated with emotional empathy (Belin et al., 2004). Neural evidence further supports this, as theta, alpha, and beta power in frontal regions increased during exposure to female voices, indicating heightened emotional engagement and relaxation (Schirmer & Escoffier, 2010).

Male hosts, while eliciting slightly less intense positive emotions overall, excel in fostering hope and confidence due to their steady tones and consistent delivery. These characteristics align with the authority principle in speech perception (Apple et al., 1979), which suggests that lower-pitched voices enhance perceived credibility and trust. EEG data from this study confirm that male voices elicit significant activation in frontal-central brain regions, associated with decision-making and motivation, which aligns with previous neuroscience findings (Johnsrude et al., 2013).

The correlation analysis further illustrates gendered differences in listener engagement, reinforcing the interactionist model of speech perception, which states that vocal characteristics dynamically interact with listener expectations and cognitive biases (Johnson, 2005). These insights suggest that podcast producers can leverage gender-specific vocal traits to optimize emotional engagement.

5.2.4 The results of the compare the positive emotions of the audience before and after listening to a podcast.

The results indicate a statistically significant enhancement in positive emotions among listeners after exposure to podcast content. This supports the mood management theory (Zillmann, 1988), which posits that media consumption, particularly through auditory stimuli, can regulate emotional states by fostering engagement and psychological well-being. Paired t-test analyses reveal notable increases in emotions

such as joy, gratitude, and passion, confirming the efficacy of podcasting as an emotionally stimulating medium.

Neurophysiological evidence further validates these findings, as EEG data demonstrate elevated delta and theta power post-listening, indicative of heightened emotional processing and attentional focus (Aftanas & Golocheikine, 2001). The observed shifts in microstate dynamics provide additional support for the theory that auditory stimuli can facilitate cognitive-emotional integration, leading to deeper listener engagement (Koenig et al., 2002). These findings align with prior research in affective neuroscience, which has established the link between auditory stimulation and enhanced emotional regulation (Schirmer & Escoffier, 2010).

Interestingly, the emotion of hope did not exhibit a statistically significant change, suggesting that certain affective states may be more resistant to auditory modulation. This finding is consistent with the cognitive appraisal theory (Lazarus, 1991), which asserts that emotions such as hope may be influenced more by an individual's personal context and situational expectations rather than external auditory input.

These results emphasize the practical implications of podcasting for emotional well-being. By carefully curating vocal delivery, content structuring, and sound design, podcast producers can maximize emotional impact and enhance listener retention. Future research should explore the longitudinal effects of podcast listening on emotional stability and psychological health, as well as potential cross-cultural variations in affective responses to auditory media.

5.3 Suggestions

5.3.1 Suggestions from the Research

Podcasts represent a rapidly growing medium for content delivery, where emotional engagement plays a crucial role in audience retention and satisfaction. Research consistently highlights the profound influence of vocal characteristics—such as pitch, rhythm, and speech rate—on listeners' emotional responses. For podcasters, optimizing these vocal elements can significantly amplify their content's emotional resonance. By doing so, they can foster deeper connections with their audience, ultimately leading to increased engagement and listener loyalty.

The modulation of pitch has been shown to evoke distinct emotional reactions. Higher vocal pitch and dynamic pitch variation, for instance, are more likely to elicit emotions such as joy and gratitude. This is especially true for female voices, which tend to be more associated with warmth and emotional expressiveness. Producers should, therefore, consider voice modulation as a strategic tool to enhance the positive emotional impact of their content. Equally, the rhythm of speech must be carefully managed; a steady rhythm may foster feelings of calm and satisfaction, whereas a more fluctuating rhythm can excite listeners and keep their attention heightened.

In terms of speech rate, a balanced approach is critical. While faster speech rates may convey enthusiasm and energy, excessively fast speech can overwhelm listeners and diminish emotional engagement. Conversely, slower speech rates may enhance clarity and emotional depth but run the risk of losing the audience's interest if they become monotonous. Podcasters need to strike a balance between these elements, ensuring that the speech rate, rhythm, and pitch align with the emotional atmosphere they aim to create.

The role of gender in emotional engagement within podcasts is another key area that content creators should consider. The current study's findings indicate that female hosts are particularly effective at evoking emotions such as love, excitement, and gratitude. This suggests that podcasts aiming to connect with audiences on a deeply emotional level may benefit from selecting female hosts or emphasizing the emotional expressiveness of their voices. However, this is not to downplay the unique emotional impact male hosts can have. Male voices were more associated with evoking feelings of hope and passion, which can be especially useful for content designed to inspire or motivate.

While these gender-based tendencies offer useful generalizations, it is essential to recognize the individual variability in vocal characteristics. Not all male or female voices will fit neatly into these patterns, and producers should focus on the specific attributes of each host's voice when designing emotionally engaging content. For instance, male hosts with softer, more melodic voices may elicit emotional responses similar to those typically associated with female voices, while female hosts with deeper tones could evoke the strength and authority more commonly linked to male voices. In

this way, voice selection becomes an intricate process that benefits from a nuanced understanding of both gendered vocal attributes and the content's emotional objectives.

Understanding the demographics of a podcast's audience is fundamental to designing content that resonates emotionally. Younger audiences, for example, tend to prefer more energetic and engaging vocal deliveries. Podcasts aimed at this demographic should focus on hosts who can maintain high levels of enthusiasm and vitality throughout the episode. This reflects broader trends in media consumption, where younger listeners seek dynamic, interactive content that keeps them constantly engaged. On the other hand, older audiences may appreciate more mature, calming vocal tones that reflect their emotional and cognitive preferences. Here, a slower pace and more deliberate vocal modulation could better serve the emotional needs of older listeners.

Producers should also consider how frequently their audience listens to podcasts. The study found that regular podcast listeners tend to form stronger emotional connections with hosts' voices, reinforcing the importance of encouraging consistent listening habits. This can be achieved through strategies such as releasing episodes on a regular schedule, promoting binge-listening of past content, and offering exclusive materials to loyal listeners. These practices can deepen the emotional bond between the host and the audience, making the podcast more than just a source of information but a consistent emotional experience that listeners look forward to.

While much of this research focused on the emotional impact of male and female voices, it is important to recognize the potential for diversity in vocal delivery to further enhance emotional engagement. Cultural background, personal experience, and even podcast genre can influence how listeners respond to various vocal characteristics. Therefore, podcast producers should not shy away from experimenting with different accents, timbres, and vocal styles. Audience feedback can be instrumental in refining these vocal strategies, helping producers identify which vocal qualities resonate most with their listeners.

Furthermore, the binary approach to gender taken in this study leaves a gap in understanding how non-binary or gender-diverse hosts may affect emotional engagement. As podcasting sharpness, it is essential to explore the impact of voices that do not conform to traditional gender categories. Non-binary voices may offer new

emotional dimensions and appeal to increasingly diverse audiences. Future research should focus on this aspect, providing a more inclusive view of how vocal characteristics influence emotional responses across a broader range of listeners.

The incorporation of electroencephalography (EEG) in understanding how vocal characteristics influence emotional engagement opens new avenues for research and application. EEG technology allows researchers to measure brainwave activity in real-time, providing direct insights into how different vocal stimuli affect listeners' emotions. By analyzing changes in specific EEG patterns, such as delta and theta rhythms, researchers can identify which vocal attributes are most effective at eliciting positive emotions.

EEG data can be used to optimize vocal delivery in various forms of auditory media, including not only podcasts but also audiobooks and voice-activated technologies. This form of neurofeedback provides a scientific basis for refining vocal techniques to enhance emotional engagement. For instance, a steady rhythm might be linked to increased delta wave activity, associated with relaxation, whereas a fluctuating pitch might enhance theta wave modulation, linked to heightened emotional arousal and engagement.

Vocal characteristics do not operate in isolation when it comes to emotional engagement. Sound design—including the use of background music, sound effects, and other auditory elements—also plays a significant role in shaping the emotional tone of a podcast. The integration of sound design with vocal delivery can create a more immersive auditory experience. For example, a host's voice accompanied by soothing background music can enhance feelings of relaxation and contentment, while upbeat sound effects may amplify feelings of joy and excitement.

Podcast producers should consider using sound design strategically to complement the emotional intent of the content. Furthermore, sound design can help create an atmosphere or emotional depth that reinforces the vocal narrative. For example, a suspenseful podcast might use eerie sound effects and a measured, deliberate vocal pace to evoke tension and anticipation.

The development of artificial intelligence (AI) and machine learning technologies has opened new possibilities for modulating vocal attributes to optimize emotional impact. AI-driven voice synthesis allows for the creation of voices that can

be tailored to evoke specific emotional responses. This could revolutionize podcast production, enabling creators to modulate pitch, tone, and rhythm dynamically in response to real-time audience feedback or pre-determined emotional objectives.

AI can also be used to create entirely synthetic voices designed to evoke specific emotional reactions, offering more control over the emotional experience of the listener. This could be particularly valuable in therapeutic or educational settings, where emotional modulation is critical to the content's success. For instance, in cognitive-behavioral therapy podcasts, AI-driven voice synthesis could be used to ensure that the vocal delivery aligns with the therapeutic goals, helping listeners achieve a state of relaxation or emotional clarity.

5.3.2 Suggestions for Future Research

The current research into podcasting, emotional engagement, and the use of vocal characteristics provides valuable insights, yet it also opens numerous avenues for future exploration. The emotional impact of podcasting, influenced by factors such as voice modulation, gender, and content type, is a complex area that warrants further in-depth examination. This section elaborates on key recommendations for future research, aiming to deepen the understanding of how auditory media affects listeners on cognitive, emotional, and physiological levels. Expanding on the initial ideas, this section will focus on genre-specific vocal characteristics, the inclusion of diverse demographic samples, neural mechanisms underpinning emotional engagement, and the investigation of negative emotions in auditory media. Each of these areas presents opportunities to refine current theories and inform more effective practices in both podcast production and broader media contexts.

While this study explored the general impact of vocal characteristics on emotional engagement, a significant gap remains regarding how these effects vary across different podcast genres. Content type likely plays a crucial role in modulating how listeners respond emotionally to vocal elements such as pitch, rhythm, and speech rate. For example, news podcasts, which often aim to convey authority and reliability, may benefit from a steady, measured vocal delivery that fosters trust and calm in listeners. In contrast, storytelling podcasts, particularly those aimed at evoking excitement or suspense, might rely more heavily on dynamic pitch variation and a fluctuating rhythm to maintain audience engagement.

Future research should aim to delineate these genre-specific effects by conducting comparative studies across a range of podcast types, including educational, comedic, and investigative genres. By doing so, researchers could better understand the optimal vocal strategies for each content type, thereby helping podcasters tailor their delivery to enhance emotional resonance with their target audience. This line of inquiry could also benefit from integrating EEG or other neuroimaging techniques to assess how brain activity varies in response to different vocal styles across genres. For instance, delta and theta wave modulation could reveal how cognitive engagement differs when listeners are exposed to educational content versus entertainment-focused podcasts.

One of the limitations of the current study is the relatively homogenous nature of the sample population, which may have constrained the generalizability of the findings. Future research should prioritize the inclusion of more diverse populations, both in terms of listener demographics and expert reviewers. Expanding the sample to include individuals from various cultural, age, and socioeconomic backgrounds would provide a more comprehensive understanding of how different groups emotionally and cognitively engage with podcasts.

Auditory processing and emotional response can be significantly influenced by demographic factors such as age, gender, and cultural background. For example, older adults may prefer slower speech rates and a more relaxed vocal tone, while younger listeners may respond more positively to fast-paced, energetic deliveries. Cultural differences in vocal preferences and emotional expression could also play a role, with some cultures valuing vocal warmth and expressiveness more than others. By incorporating a wider range of demographic variables, future studies could uncover important trends that would help podcasters tailor their content to specific audiences more effectively.

Moreover, the role of individual differences in auditory processing, such as personality traits, emotional intelligence, or sensory processing sensitivity, should be explored. These individual characteristics may modulate how listeners engage with vocal elements and the emotional content of a podcast. Researchers could employ personality assessments in conjunction with auditory stimuli to examine how traits like openness or neuroticism influence emotional responses to different vocal deliveries.

While the present research provides initial evidence of gender-specific emotional responses to podcast hosts, further investigation is required to uncover the precise neural mechanisms driving these effects. Future studies could utilize more advanced neuroimaging techniques such as functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG) to map the brain regions activated during podcast listening, particularly in relation to vocal characteristics. fMRI, with its high spatial resolution, would allow researchers to pinpoint which areas of the brain are involved in processing emotional content delivered through auditory stimuli.

This research could focus on key emotional processing centers such as the amygdala, insula, and prefrontal cortex. For example, the amygdala plays a pivotal role in the processing of both positive and negative emotions, while the prefrontal cortex is implicated in emotion regulation and decision-making. By examining the activity of these regions in response to male versus female podcasters, or high-pitched versus low-pitched voices, future research could shed light on the neural pathways that mediate emotional engagement in auditory media.

Another promising avenue involves exploring individual differences in neural responses. Some listeners may exhibit heightened emotional sensitivity to certain vocal characteristics due to variations in brain structure or function. For instance, listeners with higher levels of emotional intelligence might show increased activation in the prefrontal cortex when exposed to emotionally nuanced vocal deliveries, whereas individuals with sensory processing sensitivity may have more pronounced amygdala activation in response to fluctuating rhythms or tonal variations. Such insights could lead to personalized approaches in podcasting and other auditory media, where content is tailored to optimize emotional and cognitive engagement based on individual neural profiles.

Although this study primarily focused on positive emotional engagement, such as joy and excitement, understanding how podcasts influence a broader range of emotions—including sadness, anger, and fear—is equally important. Negative emotions, while often overlooked in media research, are a crucial part of the human emotional spectrum and can have profound effects on both cognitive and emotional processing. For example, suspenseful or investigative podcasts often aim to evoke

tension or fear, emotions that may enhance listener engagement and retention in a way that differs from positive emotions.

Future research should investigate the neural and psychological mechanisms underlying the experience of negative emotions during podcast listening. This could involve examining how different vocal characteristics, such as deeper tones or slower speech rates, contribute to the evocation of sadness or tension. Additionally, EEG and other neuroimaging methods could be used to assess how different brain wave patterns are modulated in response to negative emotional content. Understanding these mechanisms would not only expand the theoretical framework of emotional engagement in auditory media but also have practical applications in fields such as therapeutic media and mental health interventions.

For instance, podcasts that evoke controlled levels of sadness or anxiety might be used in therapeutic settings to help individuals process negative emotions in a safe and structured environment. Similarly, podcasts that evoke anger or frustration could be used as a tool for emotional regulation, helping listeners identify and manage their emotions through reflective or cathartic experiences.

Another fruitful area for future research is the intersection between vocal characteristics and sound design in shaping emotional engagement. While vocal delivery is a primary driver of emotional responses in podcasting, sound design—including background music, sound effects, and auditory pacing—also plays a significant role. The interplay between vocal and non-vocal auditory elements could be investigated to determine how these elements work together to enhance or diminish emotional engagement.

For example, future studies could explore how different types of background music influence the emotional impact of a podcast. Does upbeat, energetic music amplify the emotional intensity of an excited host, or does it serve as a distraction? Conversely, does slow, melancholic music enhance the emotional resonance of a host discussing a somber topic, or does it create cognitive dissonance These questions could be addressed using both subjective listener feedback and objective measures of brain activity to gain a deeper understanding of how sound design influences emotional processing in auditory media.

In conclusion, the future of research on podcasting and emotional engagement offers a wealth of possibilities. By expanding the focus beyond positive emotions and general vocal characteristics, future studies can explore how different genres, demographics, and neural mechanisms contribute to emotional engagement in auditory media. Advanced neuroimaging techniques, a broader exploration of emotions, and the inclusion of diverse populations will provide richer insights into how podcasts influence listeners on both emotional and cognitive levels. This knowledge can then be applied to optimize podcast production, making auditory media more effective in fostering deep emotional connections with listeners.

REFERENCES

Adaval. (2003). How Good Gets Better and Bad Gets Worse: Understanding the Impact of Affect on Evaluations of Known Brands. *Journal of Consumer Research*, 30(3), 352-367.

Alvarsson, J. J., Wiens, S., & Nilsson, M. E. (2010). Stress Recovery during Exposure to Nature Sound and Environmental Noise. *International Journal of Environmental Research and Public Health, 7*(3), 1036-1046.

Asutay, E., & Västfjäll, D. (2019). *Sound and Emotion*. *Oxford Handbook of Emotion, Social Cognition, and Everyday Problem Solving*, 23. doi: 10.1093/OXFORDHB/9780190460242.013.23

Bagozzi, R. P. (1997). Goal-directed behaviors in marketing: The role of emotion, volition, and motivation. *Psychology & Marketing*, 14(4), 309-313.

Baihua Zong. (1999). Research and analysis on the artistic style of news broadcasting. *Television Guide, 1999*(12), 92.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). *Discovering event structure in continuous narrative perception and memory*. Neuron, 95(3), 709-721. https://doi.org/10.1016/j.neuron.2017.06.041

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, 70(3), 614-636.

Bergman, P., Sköld, A., Västfjäll, D., & Fransson, N. (2021). Perceptual and emotional categorization of sound. *Journal of the Acoustical Society of America, 126*(6), 3156-3167. doi: 10.1121/1.3243297

Bidelman, G. M. (2013). The role of the auditory brainstem in processing speech and musical pitch. Hearing Research, 303, 29-38. https://doi.org/10.1016/j.heares.2013.01.010

Cai, Zuobin. (2019). Head-mounted brainwave instrument.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Transactions on Affective Computing, 1(1), 18-37.

Chenao Li. (2012). Ruminating on the Techniques and Methods of Film and Television Voiceover. *China Media Technology*, 2012(18), 64.

Chiang, J. N., Rosenberg, M. H., Bufford, C. A., Stephens, D., Lysy, A., & Monti, M. M. (2018). The language of music: Common neural codes for structured sequences in music and natural language. *Brain and Language*, *175*, 111-121.

Chion, M. (1994). *Audio-vision: Sound on screen*. Columbia University Press.

Chunyuan Han, Weixing Wang, Bojin Cheng, & Shufang Liu,. (2012). The basic problems of core strength training--core area and core stability. *Journal of Tianjin Sports Institute*, 2012, 27(02), 117-120+172.

Cole, P. M., Michel, M. K., & Teti, L. O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society for Research in Child Development*, *59*(2-3), 73-100.

Cousins, J. M. (2020). System and method for mood assessment.

Cui, X., Gao, Z., Gu, Z. (2021). Electroencephalography headband applied to depressive emotion monitoring.

Dan Peng. (2012). A comparative study on the translation of dubbed and subtitled versions of *The Sound of Music*. *Movie Literature*, *2012*(10), 167-168.

Davidson, R. J., & Fox, N. A. (1982). *Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants.* Science, 218(4578), 1235-1237. https://doi.org/10.1126/science.7146906

Dayou Ma. (2003). On indoor sound fields. *Journal of Acoustics*, 2003(2), 97-101. https://doi.org/10.15949/j.cnki.0371-0025.2003.02.001

Demos, A. P., Chaffin, R., Daniels, J. R., & Marsh, K. L. (2021). *Cognitive and neural mechanisms of passion and goal persistence: A psychophysiological perspective.* Journal of Experimental Psychology: General, 150(2), 312-328. https://doi.org/10.1037/xge0000873

Duan, S., Zhang, J., Roe, P., & Towsey, M. (2014). A survey of tagging techniques for music, speech and environmental sound. *Artificial Intelligence Review*, *41*(2), 203-221.

Duncan, Williams. (2020). Emotion in Speech, Singing, and Sound Effects. In *Advances in Human Factors and Ergonomics*, 17-26. doi: 10.1007/978-3-319-72272-6_3

Emotional Prosody. (2018). In Handbook of Clinical Neurology, 151, 293-312. https://doi.org/10.1016/B978-0-444-63247-0.00016-5

Figura, H. M., Saha, F. J., Seibt, S., Haller, H., Bringmann, H. C., Kessler, C. S., ... Jeitler, M.

(2023). Effects of an Online Meditation Course on Quality of Life and Positive Emotions: A Prospective Observational Study (EXPANSION Study). *Journal of Integrative and Complementary Medicine*.

Fois, E. (2018). 'A king of infinite space': Shakespearean Spaces in Theatre and Comics. *Between*, 8(15).

Fox, J. G., & Embrey, E. D. (1972). Music—an aid to productivity. *Applied Ergonomics, 3*(4), 202-205.

Francisco, Cuadrado., Isabel, Lopez-Cobo., Tania, Mateos-Blanco., Ana, Tajadura-Jiménez. (2020). Arousing the Sound: A Field Study on the Emotional Impact on Children of Arousing Sound design and 3D Audio Spatialization in an Audio Story. *Frontiers in Psychology*, 11:737-. doi: 10.3389/FPSYG.2020.00737

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. Philosophical Transactions of the Royal Society B: Biological Sciences, 359(1449), 1367–1377.

Fredrickson, B. L. (2013). Positive emotions broaden and build. In P. G. Devine & E. A. Plant (Eds.), Advances in Experimental Social Psychology (Vol. 47, pp. 1–53). Academic Press.

Fredrickson, B. L., & Joiner, T. (2002). Positive emotions trigger upward spirals toward emotional well-being. Psychological Science, 13(2), 172–175.

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics, 45*(3), 203-217.

Gable, S. L., & Asher, E. R. (2004). Revisiting the benefits of positive events for the positive psychology movement: The role of social comparison, positive reliving, and emotional regulation. *Behavior Therapy, 35*(2), 143-157.

Garlin, F. V., & Owen, K. (2006). Setting the tone with the tune: A meta-analytic review of the effects of background music in retail settings. *Journal of Business Research, 59*(6), 755-764.

Gary, E., Strangman, V., Ivkovic, G., Zhang, Q. (2020). Wearable brain imaging with multimodal physiological monitoring. *Journal of Applied Physiology, 124*(3), 564-572.

Gaver, W. W., & Mandler, G. (1987). Play it again, Sam: On liking music. *Cognition and Emotion, 1*(3), 259-282.

Gibbs, T. (2007). *The Fundamentals of Sonic Art and Sound Design*. Ava Publishing.

Giftakis, J. E., Rise, M. T., Carlson, D. L., Stypulkowski, P. H., Stanslaski, S. R., Jensen, R. M., Denison, T. J. (2022). Mood circuit monitoring to control therapy delivery

Gonzalo, Iturregui-Gallardo., Anna, Matamala. (2021). Audio subtitling: voiceover and voice-over effects and their impact on user experience. *Perspectives-studies in Translatology*, 29(1), 64-83. doi: 10.1080/0907676X.2019.1702065

Gorbman, C. (1987). Unheard melodies: Narrative film music. Indiana University Press.

Gordon, R. L., Jacobs, M. S., Schuele, C. M., & McAuley, J. D. (2021). PRISM: A unified framework of rhythm's role in speech and music perception. Nature Reviews Neuroscience, 22(2), 98–113. https://doi.org/10.1038/S41583-020-00371-X

Groeneweg, G., Stan, E. A., Celser, A., MacBeth, L., & Vrbancic, M. I. (1988). The effect of background music on the vocational behavior of mentally handicapped adults. *Journal of Music Therapy, 25*(3), 118-134.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271-299.

Guéguen, N., Jacob, C., Le Guellec, H., Morineau, T., & Lourel, M. (2008). Sound level of environmental music and drinking behavior: A field experiment with beer drinkers. *Alcoholism: Clinical and Experimental Research, 32*(10), 1795-1798.

He, W. (2019). Bertolt Brecht's theatrical concept of alienation effect and the Chinese application and transformation. *Journal Title Omitted*, 46(1).

Hossan, A., & Chowdhury, A. M. (2016, May). Real-time EEG based automatic brainwave regulation by music. *In 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)* (pp. 780-784). IEEE.

Howard, D. J., & Gengler, C. (Year). Emotional Contagion Effects on Product.

Irwin, A., Hall, D. A., Peters, A., & Plack, C. J. (2011). Listening to urban soundscapes: Physiological validity of perceptual dimensions. *Psychophysiology, 48*(3), 325-337.

Jeliruew. (2013). The evolution of sound recording technology: nuance and specialization in the 1930s and 1940s. *Sound Recording and Technology Review*, Volume (if available), pp.

Jensen, P. R. (2020). The future of education [Audio podcast episode]. In A. Johnson (Ed.), *Insights in Education.* Academic Podcast Network.

Jin, J., Shi, J., & Zhang, G. (2019). Application of Medical Feature Extraction Based on Image Technology. In *Proceedings of the 2nd International Conference on Robotics, Control & Automation Engineering (RCAE 2019)*.

Jung-Ki, Nam., Eungsu, Kim., Jieun, Kim., Jaeyun, Kim., Daeho, Lee., Won, Young, Chung. (2023). Effect of audio content on simple and repetitive tasks requiring eyesight: Focus on task performance and emotional responses. *Information Development*, *026666692311692-026666692311692*. doi: 10.1177/02666669231169242

Junying Li. (2000). Sound Modeling Art of Translated Films. *Exploration and Research on Translated Films, China Radio, Film and Television Publishing House*, 2000(08), 575.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? Psychological Bulletin, 129(5), 770-814.

Juslin, P. N., & Laukka, P. (2003). Emotional expression in speech and music: Evidence of cross-modal similarities. Annals of the New York Academy of Sciences, 1000(1), 279-282. https://pubmed.ncbi.nlm.nih.gov/14766637/

Juslin, P. N., & Sloboda, J. A. (2010). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. Behavioral and Brain Sciences, 31(5), 559–621. https://doi.org/10.1017/S0140525X08005293

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559-575.

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2020). Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, 2020.

Kassabian, A. (2001). Hearing film: Tracking identifications in contemporary Hollywood film music. Taylor & Francis.

Kathryn, Sharpe, Wessling. (2019). Positive Design: A Positive Psychology Workout for

Design Thinkers.

Ke Jin. (2015). The Road of Idolization of Film and Television Voiceover Artists. *Western Radio and Television*, 2015(13), 106-107.

Kelly Reynolds. (2020). *Emoacoustics: A Study of the Psychoacoustical and Psychological Dimensions of Emotional Sound Design*. *Journal of The Audio Engineering Society, 60*, 21-28

Killin, A., & Killin, A. (2020). *The Polysemy Theory of Sound*. *Erkenntnis*, 1-24. doi: 10.1007/S10670-019-00201-7

Klimesch, W. (1999). *EEG alpha and theta oscillations reflect cognitive and memory performance: A review*. Neuroscience & Biobehavioral Reviews, 22(3), 427-436. https://doi.org/10.1016/S0149-7634(98)00056-3

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2020). *EEG alpha and theta oscillations reflect cognitive and emotional processing: A review*. Neuroscience & Biobehavioral Reviews, 118, 165-180. https://doi.org/10.1016 /j.neubiorev.2020.07.013

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. Psychology of Music, 43(2), 155-170.

Kreiman, J., & Sidtis, D. (2011). Foundations of voice studies: An interdisciplinary approach to voice production and perception. Wiley-Blackwell.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. Psychological Bulletin, 126(1), 159–179. https://doi.org/10.1037/0033-2909.126.1.159

Ladd, D. R. (2008). Intonational phonology. Cambridge University Press.

Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). *Electroencephalographic signatures of attentional and cognitive control in the resting brain*. The Journal of Neuroscience, 23(27), 9097-9106. https://doi.org/10.1523/JNEUROSCI.23-27-09097.2003

Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. Cognition and Emotion, 19(5), 633–653. https://doi.org/10.1080/02699930441000445

Lee, A. Y., & Sternthal, B. (1999). The effects of positive emotions on memory. *Journal of Consumer Research, 26*(2), 115-127. https://doi.org/10.1086/209554

León, J., Moya-Albiol, L., & Martínez-Selva, J. M. (2021). *Empathy and EEG alpha*

asymmetry in different emotional contexts: The role of social bonding. Brain and Cognition, 145, 105633. https://doi.org/10.1016/j.bandc.2021.105633

Lewis, J. (2009). Human Neuroemotional Research. *The Journal of Neuroscience, 29*(7), 2283-2296.

Li, M., Wang, H., Zheng, Z., & Wang, Y. (2015). The role of sound in enhancing immersion and affective responses in computer games. *International Journal of Human-Computer Interaction, 31*(8), 553-564.

Liang Wang. (2023). A cognitive study of Sound design and emotional resonance in acoustic spaces. *Art and Design*, 8.

Lietz, R., Harraghy, M., Calderon, D., Brady, J., Becker, E., Makedon, F. (2019). Survey of mood detection through various input modes, 28-31. doi: 10.1145/3316782.3321543

Lihong Hou. (1998). Artistic Styles of Film and Television Voiceover from the Cultural Differences between the East and the West. *Chinese Television*, 1998(01), 41-44.

Lihong Li. (2007). Re-creation of Acoustic Language in Performance Art - An Analysis of Voiceover in Film and TV Drama. *Contemporary Film*, 2007(06), 95-98.

Lihong Li. (2010). Voiceover in Translated Films. *Contemporary Cinema*, 2010(05).

Liu, R., Wang, Y., Shu, M., & Gao, T. (2019). An algorithm study of electrocardiogram signal denoising by using wavelet transform method. In R. Liu, Y. Wang, M. Shu, & T. Gao (Eds.), *2019 2nd International Conference on Computer Information Science and Application Technology (CISAT 2019)*.

Liu, Z., Zhang, H. (2019). Mood monitoring device, system, and method.

Luce, M. F., Payne, J. W., & Bettman, J. R. (1999). Emotional Trade-Off Difficulty and Choice. *Journal of Marketing Research*, 36(2), 143-159.

Lutz, A., McFarlin, D. R., Ricard, M., & Davidson, R. J. (2020). *Gamma synchrony and goal-directed behavior: Neural mechanisms of hope and perseverance.* Proceedings of the National Academy of Sciences, 117(8), 4457-4465. https://doi.org/10.1073/pnas.1918970117

MA, Chun-dong. (2022). Analysis of Cognitive Psychological Properties in Soundof Products. *Packaging Engineering*.

Ma, W., & Thompson, W. F. (2015). Human emotions track changes in the acoustic environment. *Proceedings of the National Academy of Sciences of the United States of*

America, *112*(47), 14563-14568.

Macalister, R., Moorhouse, S. S., Cooke, T., Evans, L. (2019). A mood indicator.

Maria, T., & Russo, T. (2010). Communication (action with communicative content). *Clinica Terapeutica*, *161*(2), 193-194.

Matthias, S., Aslan, A., & Bach, D. R. (2020). Primary auditory cortex representation of fear-conditioned musical sounds. *Human Brain Mapping*, 2020.

Medvedev, O., Shepherd, D., & Hautus, M. J. (2015). The restorative potential of soundscapes: A physiological investigation. *Applied Acoustics*, *96*, 20-26.

Mei Han. (2017). Historical and Practical Reflections on the Vocal Theory of "Dantian Breath". *Modern Communication (Journal of Communication University of China)*, 39(07), 96-98.

Mikko, Viinikainen., Jari, Kätsyri., Mikko, Sams. (2022). Representation of perceived sound valence in the human brain. *Human Brain Mapping*, 33(10), 2295-2305. doi: 10.1002/HBM.21362

Miller, R. G. Jr. (1997). *Beyond ANOVA: Basics of Applied Statistics.* CRC Press.

Milliman, R. E. (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research*, *13*(2), 286-289.

Minggang Chen. (2014). Analysis on the innovation of language art of radio broadcast hosting. *Communication Research*, *2014*(2), 138.

Moll, J., Zahn, R., & Grafman, J. (2019). *The neuroscience of gratitude and its impact on social behavior*. Nature Reviews Neuroscience, 20(5), 293-305. https://doi.org/10.1038/s41583-019-0157-4

Morris, J. D., Woo, C., Geason, J. A., & Kim, J. (2002). The Power of Affect: Predicting Intention. *Journal of Advertising Research*, 42(3), 7-17.

Niklas, Rönnberg., Jonas, Löwgren. (2021). The sound challenge to visualization design research. In *European Conference on Product and Process Modeling*, 31-34. doi: 10.3384/ECP10305

Pannese, A. (2012). A gray matter of taste: Sound perception, music cognition, and Baumgarten's aesthetics. *Studies in History and Philosophy of Biological & Biomedical Sciences*, 43(3).

Papotti, D. (2013). Luoghi, territori e paesaggi del teatro: per un approccio geografico al

rapporto fra azione teatrale e dimensione spaziale. *Ricerche di S/Confine*, Dossier 1.

Piazza, E. A., Theunissen, F. E., Wessel, D., & Whitney, D. (2020). Publisher Correction: Rapid Adaptation to the Timbre of Natural Sounds. *Scientific Reports*, 2020.

Prosody (Linguistics). (2023). In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Prosody_(linguistics)

Qin, X., Sheng, Y., Liu, S., & Chai, L. (2022). Graph Wavelet Transform-based Denoising for Dynamic PET Images with Multi-Feature Information. In X. Qin, Y. Sheng, S. Liu, & L. Chai (Eds.), *Proceedings of the 41st China Control Conference*.

Qin, Y., Bulbul, T. (2023). Electroencephalogram-based mental workload prediction for using Augmented Reality head-mounted display in construction assembly: A deep learning approach. *Automation in Construction, 152*, 104892. doi: 10.1016/j.autcon.2023.104892

Qing Miao. (2000). Discussion on the authenticity of sound in TV dramas. *China Television*, *2000*(S1), 53-55.

Quoidbach, J., & Mikolajczak, M. (2010). Positive interventions: An emotion regulation perspective. *Psychological Bulletin, 136*(5), 655-694.

Rahman, J. S., Gedeon, T., Caldwell, S., & Jones, R. (2020, July). Brain melody informatics: analysing effects of music on brainwave patterns. *In 2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

Ratcliffe, E., Gatersleben, B., & Sowden, P. T. (2013). Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology, 36*, 221-228.

Russell, J. A. (1980). A cyclical model of emotion. *Journal of Personality and Social Psychology*, *39*(6), 1161.

Sachs, M. E., Habibi, A., Damasio, A., & Kaplan, J. T. (2020). *Dynamic intersubject neural synchronization reflects affective responses to sad music.* NeuroImage, 206, 116316. https://doi.org/10.1016/j.neuroimage.2019.116316

Sascha, Frühholz., Wiebke, Trost., Sonja, A., Kotz. (2020). The sound of emotions-Towards a unifying neural network perspective of affective sound processing. *Neuroscience & Biobehavioral Reviews*, 68, 96-110. doi: 10.1016/J.NEUBIOREV.2016.05.002

Sawka, M. N., Friedl, K. E. (2022). Emerging Wearable Physiological Monitoring

Technologies and Decision Aids for Health and Performance. *Journal of Applied Physiology*, *124*(2), 430-431.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. Speech Communication, 40(1-2), 227-256.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. Speech Communication, 40(1-2), 227-256. https://doi.org/10.1016/S0167-6393(02)00084-5.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33*(3), 239-251.

Science - Linguistics; Researchers from University of Piemonte Orientale Detail New Studies and Findings in the Area of Linguistics (Comparing insults across languages in films: Voiceover as cross-cultural mediation). *Science Letter*, 2019.

Seligman, M. E. P. (2002). Positive psychology, positive prevention, and positive therapy. *Handbook of Positive Psychology*, 2(2), 3-12.

Seligman, M. E. P. (2002). Positive psychology, positive prevention, and positive treatment. *Handbook of Positive Psychology, 2*(2), 3-12.

Sharda, M., & Singh, N. C. (2012). Auditory perception of natural sound categories – An fMRI study. *Neuroscience, 220*, 203-212.

Shen, Y. C., & Chen, T. C. (2006). When East meets West: The effect of cultural tone congruity in ad music and message on consumer ad memory and attitude. *International Journal of Advertising, 25*(1), 51-70.

Shuguang Peng. (2014). Exploring the language characteristics and skills of radio broadcast hosting. *Media Forum, 2014*(2), 72-73.

Shuhua Zhang. (2002). Advances in interpersonal communication skills research. *Psychological Science, 25*(4), 503-505.

Sifan, Wu., Fei, Li., Yuan, Yao. (2023). Deep Learning for the Detection of Emotion in Human Speech: The Impact of Audio Sample Duration and English versus Italian Languages. *IEEE Wireless Communications and Networking Conference (WOCC)*, 1-6. doi: 10.1109/WOCC58016.2023.10139686

Smith, C. R. (2015). Principles and practice of voiceover foreign-language films. Springer.

Smith, S. M. (1985). Background music and context-dependent memory. *The American Journal of Psychology*, *591-603.*

Smith, S. M., & Taylor, J. R. (2022). Neural correlates of storytelling and emotional *engagement: Insights from EEG studies*. Journal of Cognitive Neuroscience, 34(5), 782-796. https://doi.org/10.1162/jocn_a_01849

Sogand, Mohammadi., M., Gouran, Orimi., Hamid, R., Rabiee. (2023). Enhancing the Prediction of Emotional Experience in Movies using Deep Neural Networks: The Significance of Audio and Language. *arXiv.org, abs/2306.10397*. doi: 10.48550/arXiv.2306.10397

Song Zhang. (1985). On the tone of broadcasting. *Modern Communication-Journal of Communication University of China, 1985*(1), 62-67+24. doi:10.19997/j.cnki.xdcb.1985.01.016.

Song Zhang. (1989). On the creative method and expression law of broadcasting. *Chinese Journal of Radio and Television, 1989*(6), 44-49.

Tan, S.-L., Pfordresher, P. Q., & Harré, R. (2020). *Psychology of Sound: From Sound to Significance* (Second Edition). doi: 10.4324/9781315648026

Temporal Envelope and Fine Structure. (2023). Acoustical Science and Technology, 44(2), 121-137. https://doi.org/10.1250/ast.44.121

Thompson, W. F. (2015). *Music, Thought, and Feeling: Understanding the Psychology of Music*. Oxford University Press.

Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. The Journal of the Acoustical Society of America, 134(1), 628-639. https://doi.org/10.1121/1.4807565

Ting, Y. (2015). The Research on the Impact of Background Music on Appetite of Customers in Restaurant. *The Open Cybernetics & Systemics Journal*, 9(1).

Västfjäll, D., Juslin, P. N., & Hartig, T. (2012). Music, subjective well-being, and health: The role of everyday emotions. In *Music and the Social Sciences: Theories, Research, and Applications* (pp. 403-425). Psychology Press.

Wagner, M. J. (1975). Effect of music and biofeedback on alpha brainwave rhythms and attentiveness. *Journal of Research in Music Education, 23*(1), 3-13.

Walker, J. L. (1977). Subjective reactions to music and brainwave rhythms. *Physiological*

Psychology, 5(4), 483-489.

Wang, G., Wang, F.-F., Xiong, F., & Zhang, Z.-R. (2022). Pulse Signal Preprocessing Based on Wavelet Transform-Principal Component Analysis. In G. Wang, F.-F. Wang, F. Xiong, & Z.-R. Zhang (Eds.), *The 15th Textile Bioengineering and Informatics Symposium*.

Wang, Y.-H. (2014). Rethinking Hanslick: Music, Formalism, and Expression. *Music Library Association. Notes*, 71(2).

Watzlawick, P., Beavin, J., Bavelas, B., & Jackson, D. D. (1975). Pragmatics of human communication. *Journal of Communication, 27*(1), 143-. doi:10.2307/3206369

Wei Hu. (2018). Analyzing the Application of Lines in Creating Character Image. *House of Drama*, 2018(22), 20-21.

Woolley, S. M. N., Fremouw, T. E., Hsu, A., & Theunissen, F. E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nature Neuroscience, 8*(10), 1371-1379.

Wu Yu. (1999). The skills and characteristics of radio broadcasting hosting in the field of new media. *Journalist's Cradle, 1999*(4), 137-138.

Wujiang Zhang. (2013). The application of translation creation in game localization translation. *Modern Communication (Journal of Communication University of China), 2013, 35*(12), 165-166.

X, Zhang., & A, Yin. (2017). Research on the way of measuring music-evoked emotions. *Journal of Central Conservatory of Music*, 2017(1), 14.

Xiang, Y., Guo, J., Chen, Y., Yin, K., Sun, S., & Han, C. (2023). Heart rate estimation algorithm for radar signals based on wavelet threshold noise reduction and Resnet50. In Y. Xiang, J. Guo, Y. Chen, K. Yin, S. Sun, & C. Han (Eds.), *Proceedings of the 35th China Control and Decision Making Conference*.

Xingchen Dong. (2018). Talking about the artistic creation and diversified expression of film and television voiceover. *Science and Technology Communication*, 10(13), 69-70.

Xinhua News Agency . (2023). White Paper on National Listening Trends. Beijing: State Key Laboratory of Media Convergence Production Technology and System, Xinhua News Agency.

Xiting Huang.(2007). Neural processing of psychology. *Nature Reviews Neuroscience, 15*(6),

355-366.

Xueren Xu. (2023). Changes and Trends of Film and Television Voiceover under the Iteration of Recording Technology. *East China Normal University*.https://link.cnki.net/doi/10.27149/d.cnki.ghdsu.2023.001765 doi:10.27149/d.cnki.ghdsu.2023.001765

Xueyi Ma. (2007). Life and Emotion Created by Sound--Talking about the Shaping of Animal Characters by Voiceover in American Animated Films in Recent Years. *Movie Review*, 2007(23), 18-19.

Yang, Y., Shillingford, B., Assael, Y., Wang, M., Liu, W., Chen, Y., ... & de Freitas, N. (2020). Large-scale multilingual audio visual voiceover. *arXiv preprint arXiv:2011.03530*.

Yi-Yoen, Kim., Jun-Woo, Kim., Min-Woo, Lee., Gyoung-Sun, Kim., Jae-Sang, Cha., Kyung-Hoon, Kim. (2022). Sound source matching LED emotion contents design. In *2015 7th International Conference on Ubiquitous and Future Networks (ICUFN)*, 372-374. doi: 10.1109/ICUFN.2015.7182567

Yuan Wang. (2014). Characteristics and skills of broadcasting hosting in media vision. *Journalist Cradle*, *2014*(4), 93-94.

Yuanyuan Feng. (2013). The function of character voice modeling in film and television voiceover. *News Knowledge*, 2013(06), 61-62+81.

Yue Jia. (2015). On the "Fitness" of Characterization in Film and Television Voiceover. *China Out-of-School Education*, 2015(20), 22-23+19.

Zhuang, T., Zhao, H., & Tang, Z. (2009). A study of brainwave entrainment based on EEG brain dynamics. *Computational Information Science, 2*(2), 80-86.

BIOGRAPHY

NAME	Jun Ji
DATE OF BIRTH	October 03, 1986
PLACE OF BIRTH	China
ADDRESS	Room 2301, Unit 2, Building 2,Meiqihua Community, No.279,Sheng li West Street, Xinghualing District, Tai Yuan City, Shanxi Province China
POSITION	Lecturer
EDUCATION	2005-2009 Bachelor of Arts, Chengdu University of Technology 2015-2018 Master of Fine Arts, School of Art, Huazhong University of Science and Technology 2022-2025 (Ph.D.) Doctor of Philosophy in Creative Media, Faculty of Information Sciences, Mahasarakham University
Research output	Ji, J. (2021). Design and implementation of online teaching and assessment models in the context of "Internet +": Taking "evaluator of etiquette hosts" as an example. Science and Technology Communication, 08, 141-143. https://doi.org/10.16607/j.cnki.1674-6708.2021.08.048 Ji, J. (2018). An exploration of curriculum reform in college media majors in the era of all media: A study of curriculum reform in film and television dubbing for broadcasting and hosting majors in Shanxi Province. Media Forum, 15, 21-22. https://doi.org/CNKI:SUN:CMLT.0.2018-15-012 Ji, J. (2017). A brief discussion on the coordination of guqin performance and ancient poetry recitation. Drama Home, 18, 76. https://doi.org/CNKI:SUN:XJZT.0.2017-18-051 Ji, J. (2014). Research on practical teaching of film and television majors. Art Technology, 05, 31. https://doi.org/CNKI:SUN:YSKK.0.2014-05-035 Ji, J. (2020). The impact and value reconstruction of AI technology on the skill requirements of film and television dubbing talents. Film Studies, 01, 36-37. https://doi.org/CNKI:SUN:IYYJ.0.2020-01-012 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

Jun, Ji., Youngmee, K., Liamthaisong, K. (2025). Emotional Responses in Podcast Sound Design: A Frequency-Based Analysis of the Host's Voice Spectrum. Journal of Communication Arts Review, 29(1); January-April , 82-92. (TCI Q1)

Jun Ji., Kotchaphan Youngmee, and Khachakrit Liamthaisong. (2025) Integrating Acoustic Modulation and Listener Demographics for Enhanced Podcast Emotional Resonance. Operationnal Research in Engineering Sciences : Theory and Applications (ORESTA). 8(1) ; 13 Feb. (Scopus Q1)