



Emotional Responses in Podcast Sound Design: A Frequency-Based Analysis of the Host's Voice Spectrum

Jun Ji. *Faculty of Informatics, Mahasarakham University.*

Kotchaphan Youngmee. *Faculty of Informatics, Mahasarakham University.*

Khachakrit Liamthaisong. *Faculty of Informatics, Mahasarakham University.*

E-Mail: 51431046@qq.com, kotchaphan@msu.ac.th, khachakrit.l@msu.ac.th

Received: 27 July 2024; Revised: 30 July 2024; Accepted: 16 October 2024

Abstract

Sound design is essential in digital media for influencing listeners' perceptions of sound and their emotions, particularly in podcasting. This study looks at how podcast hosts' voice images are affected by acoustic spectrum analysis. We demonstrate that the voice of the host of the Happy Planet podcast exhibits severe oscillations in its high-frequency components, which are primarily above 2 kHz with a core frequency surpassing 3 kHz. Emotional engagement and clarity are improved by this spectral profile. On the other hand, the mid-frequency spectrum is smoother and adds to a pleasing aural experience. It spans from 500 Hz to 2 kHz, with a center frequency of approximately 1 kHz.

According to our research, mid-frequency components offer comfort and efficient emotional transmission, while high-frequency components are essential for enhancing sound recognition and emotional expression. Optimized high-frequency spectra lead to a 22% increase in listener retention and a 30% increase in emotional connection, according to quantitative studies. Gender disparities were noted: male listeners preferred sounds with a lot of bass, whereas female listeners preferred a balance of high and mid-frequency components. This study demonstrates how frequency modulation can modulate emotional impact and listener involvement, supporting Ma and Thompson's (2015) approach that links sound qualities to psychological responses. Our findings highlight the value of customized sound design in podcast production and provide guidance for improving emotional resonance and listener experiences. This study highlights the potential of sound design to enhance digital media content while also advancing the theoretical knowledge of its role in media and communication.

Keywords: Sound Design, Acoustic Analysis, Emotional Response, Podcast, Media Studies

Introduction

In the digital media era, sound design is crucial in film, television, games, virtual reality, and podcasts. Sound not only conveys information but also stimulates emotions. Research has demonstrated that sound design significantly impacts viewers' emotional experiences and cognitive processes (Bradley & Lang, 2019). Podcasts have emerged as a popular medium for content consumption, making sound design in this domain particularly critical. Understanding the emotional impact of sound design in podcasts is essential, as it directly influences audience engagement and satisfaction.

The theoretical foundations of the relationship between sound design and positive audience affective responses are rooted in the interdisciplinary fields of psychoacoustics, cognitive psychology, and media studies. However, although there have been a number of studies that have explored the impact of sound design on audience affective responses, theoretical research on its specific mechanisms and effects in terms of positive audience affective responses is still relatively limited (Koelsch, 2021).

It is commonly known that hearing is a crucial sense in the human experience and that processing emotional cues depends on it. Certain acoustic characteristics, including as pitch, rhythm, and timbre, have been linked to a range of emotional reactions, according to psychoacoustic research. For instance, high pitches are frequently linked to tension or excitement, but low frequencies can evoke feelings of fear or relaxation (Lerner et al., 2016). certain results imply that sound design might purposefully alter certain factors to elicit particular emotional responses from listeners. With a particular focus on podcast sound design, this study seeks to close the knowledge gap on the important role sound plays in evoking emotional reactions. In podcasts, as opposed to other media, the auditory sense serves as the main means of communication, which emphasizes the significance of precisely manipulating acoustic qualities.

Hearing is vital in processing emotional cues, with psychoacoustic studies identifying parameters like pitch, rhythm, and timbre that trigger different emotions. For example, high pitches are associated with excitement, while low-frequency sounds may induce calm or fear (Lerner et al., 2016). These findings suggest that sound design can intentionally manipulate these parameters to elicit specific emotional states.

Cognitive psychology examines how individuals interpret and process auditory information within their cognitive schema and emotional memory. Emotional resonance theory and the mirror neuron system highlight humans' innate ability to empathize with emotions expressed by others, including through auditory means (Scherer, 2017). Thus, sound design can exploit these cognitive processes to create immersive and emotionally resonant experiences.

Media studies have further extended this theoretical field by examining how cultural and contextual factors shape interpretations of sound design. The semiotics of sound is a key area that involves the study of the meaning of sound in cultural contexts. The cultural connotations of a particular sound or musical theme can significantly influence an audience's emotional response, and the reception of a media work in a wider socio-cultural context can also be affected (Zatorre & Salimpoor, 2018).

Significance of the study:

This study is significant because it may offer further information about how to optimize podcast sound design to elicit favorable emotional reactions. The results of this investigation on the particular acoustic characteristics that affect listeners' emotions may help the podcasting business create sound design tactics that are more successful.

This research will contribute to our understanding of the emotional influence of sound design while also providing useful guidance for podcasters and content makers by bridging the theory-practice divide. This study has ramifications that go beyond scholarly discourse. It may have an impact on how sound design is used in various mediums.

Research objectives:

Through the research in this paper, it is expected that the following objectives will be achieved:

- (1) To explore the role of frequency spectrum in podcast host voice design.
- (2) To evaluate the impact of spectral features on listeners' positive emotional responses.
- (3) To investigate practical applications of voice spectrum design in enhancing listener engagement.

Research Contributions

(1) In-depth Exploration of Voice Spectrum in Podcasting. This study offers a comprehensive examination of how the frequency spectrum of a podcast host's voice influences listeners' emotional responses. By dissecting various spectral elements—such as pitch, tone, and frequency variations—it provides a nuanced understanding of how these factors shape the auditory experience. This research enhances existing knowledge by creating a precise framework that aids in designing podcast content that effectively connects with listeners on an emotional level, addressing a gap in the literature related to the application of voice spectrum in podcast production.

(2) Practical Sound Design Strategies for Emotional Engagement. The findings of this study offer actionable insights for sound designers and content creators, particularly within the podcasting domain. By demonstrating how specific spectral characteristics—like the interplay between different frequency bands and the dynamic modulation of the voice—can be leveraged to evoke targeted emotional reactions, this research contributes to the development of sound design techniques that are both technically sophisticated and emotionally resonant. These strategies are aimed at producing podcast content that engages and emotionally connects with audiences, ultimately enhancing listener experience and loyalty.

(3) Empirical Support for the Impact of Spectral Voice Design. Through rigorous experimentation and data analysis, this research substantiates the claim that the spectral design of a podcast host's voice plays a significant role in shaping listener emotions. The study provides empirical evidence that supports the effectiveness of specific spectral attributes in conveying emotions, offering a scientific foundation for these sound design techniques. This validation reinforces the theoretical principles of voice spectrum design and serves as a benchmark for future research within audio engineering, media psychology, and communication studies.

2. The use of physical properties of sound in sound design

2.1 Correspondence between physical properties of sound and psychological parameters

Sound's physical properties, such as frequency, intensity, waveform, and time, correspond to psychological parameters like pitch, loudness, timbre, and duration. These relationships have been extensively explored in music psychology. By controlling sound's physical properties, one can influence the corresponding psychological parameters of the audience, thereby aligning sound effects with the intended design to enhance the work's impact and immersion (Collins, 2020). An individual's voice characteristics, including timbre and tone, are shaped by factors such as upbringing, pronunciation habits, and vocalization style. Timbre reflects the unique qualities of the sound source.

2.2 Style characteristics of sound design

In film, television, and video games, sound design is often integrated with other artistic styles to create a cohesive experience. Sound designers achieve stylistic unity by controlling and adjusting the physical properties of sound.

In the mobile game project "Baolian, the sound design included unifying the sound effects into short samples of approximately 0.8 seconds, minimising the use of low-frequency sounds in order to maintain a bright and fresh tone. Instruments such as wind chimes and marimba were used, and short, fast major scale rises or major third arpeggio rises were employed. These design choices are intended to create an emotional atmosphere of stability, relaxation and pleasure. These stylistic features reflect everyday listening experiences and aesthetic perceptions (Gibson & Polfreman, 2021). Generally speaking, the bass is turbid and heavy, the alto is mellow, and the treble is light. These responses are built on the basis of the general people's life experience and aesthetic experience, and are generated through association and imagination.

Sound design in film, television, and games must align with the intended emotion and meaning of the work. Over-exaggerated sound in a realistic style can be jarring. Sound designers should ensure that the emotional tone is consistent by carefully controlling the physical properties of sound.

2.3 Application of auditory cognition effect

Auditory cognition emerges from the sensory processing of sound by the human auditory system, which is influenced by individual life experiences, emotional states, cultural context, and specific cultural backgrounds (Tse et al., 2022). Understanding these cognitive processes is crucial for sound design, as it allows for the alignment of auditory stimuli with the psychological and emotional responses of the audience.

Sound designers use both the physical characteristics of sound and the acuity of the human ear to create a realistic soundscape. For example, using sound attenuation properties to alter how sound is perceived spatially results in high-frequency loss and volume decrease for distant sounds since high frequencies attenuate more quickly than low frequencies.

Since the human ear is particularly sensitive to the middle frequencies (500–1000 Hz), which improves pitch discrimination and sound perception, amplifying these frequencies works well to give the sensation of proximity. Furthermore, it's crucial to take into account variations in the overall spectrum balance and sound intensity in addition to the pitch shift when applying the Doppler effect to noises like racing vehicles or airplanes.

3. Application of sound aesthetic psychology in sound design

In the practice of sound design, the choice of sound effects and sound is crucial, and music as a form of sound, in which also plays an important role, the two are intertwined, and together they create a more realistic contextual experience for the listener.

3.1 The association between sound form elements and emotional experience

Human emotion is bipolar, including the increase or decrease of emotional power, the fluctuation of the degree of excitement and the change of emotional intensity. There is an obvious correspondence between this bipolarity and sound form elements, such as the height of the pitch, the harmony and dissonance of the tone, the change of the stability of the sound, the speed of the rhythm and the length of the rhythm as well as the volume. Based on the theory of “heterogeneity and homogeneity” in Gestalt psychology, there is an intrinsic connection between the formal elements of sound and human emotions, which leads to similar psychological and emotional experiences.

3.2 Sound form elements and cultural context

The cultural context of sound aesthetics needs to be considered when designing podcasts. The host's voice must resonate with the audience's cultural and emotional expectations to create a compelling listening experience. For example, in a podcast aimed at a Western audience, employing a voice modulation that aligns with Western musical traditions can enhance listener connection. Utilizing specific scales or rhythms can evoke familiar emotional responses, thereby enhancing the listener's engagement. The orchestration of the host's voice, through careful selection of pitch, tempo, and modulation, can create a desired atmosphere, whether it is calming, authoritative, or exciting. By adhering to these principles, the sound designer ensures that the podcast delivers a consistent and emotionally resonant auditory experience.

3.3 The key role of sound design in positive mood shaping

Recent research emphasizes the importance of sound in creating emotional reactions. According to Stephen P. Pliva's research, music can elicit emotions such as excitement and stress, which are connected to brain regions responsible for emotional processing (Sek & Moore, 2020). In podcast sound design, the frequency of the host's voice is important for influencing listener emotions. Frequency, measured in Hertz (Hz), influences pitch; higher frequencies frequently indicate excitement or haste, whilst lower frequencies evoke calmness or authority. Sound designers can improve emotional involvement and the auditory experience by strategically manipulating frequency.

Research methodology

This paper adopts a mixed research methodology, combining qualitative and quantitative research strategies in order to fully explore the research questions. Firstly, quantitative data was collected through a structured questionnaire designed to cover the key variables in the study and statistical analysis to reveal trends and patterns in the data. Secondly, qualitative data were obtained through semi-structured in-depth interviews, which were designed to provide an in-depth understanding of respondents' subjective experiences and perspectives in order to provide contextual support and interpretation of the quantitative data. To ensure the reliability of the research instrument, this study will validate the process of using, collecting and analysing the questionnaire and interview data, focusing on checking the internal consistency of the questionnaire and the consistency of the content of the interviews in order to ensure the accuracy and reliability of the data.

Qualitative Research

In this study, a qualitative analytical approach is used to provide insights into the use of audio spectral design in podcast sound and its impact on listeners' affective responses. This approach aims to reveal the intrinsic mechanisms behind sound design and its specific impact on listener experience. We used semi-structured in-depth interviews with respondents from different backgrounds to gain insights into audio design and its impact on listener emotions. Specifically, we chose the following three categories of respondents and conducted a total of 2 interviews for each category, totalling 30 respondents:

1. Audio designers: 10 audio designers with extensive sound design experience were selected who were able to provide insights into how spectral features can be used in sound design and discuss how these features can enhance emotional expression.

2. Podcast Producers: 10 active podcast producers who have been directly involved in content creation and sound design were selected to share the application of spectral design in real-world creation and its impact on listener feedback.

3. Regular podcast listeners: 10 regular podcast listeners were selected who were able to provide a direct experience of the podcast's sound and feedback on the practical effects of the sound design on their emotional response.

Quantitative Research

In this study, the quantitative analyses focused on a systematic and quantitative assessment of spectral features in podcast sound design. By designing and distributing questionnaires, we collected listeners' subjective responses to different sound designs. Subsequently, we statistically analysed these subjective feedbacks with actual sound spectral data to reveal how spectral features affect listeners' emotional responses. Through in-depth analysis of the questionnaire data, we were able to scientifically assess the effectiveness of sound design in conveying emotion, and thus provide targeted optimisation recommendations for podcast sound design.

Research results

In sound design, the analysis of the high-frequency sound spectrum is crucial in shaping the voice image of a podcast host. The high-frequency sound spectrum often reflects the sharpness, clarity, and activity of the voice, which directly affects the listeners' perception and emotional experience of the host's voice. The following is frequency sound spectrum analysis of the presenter's voice in the Happy Planet programme on the Himalayan podcast platform:

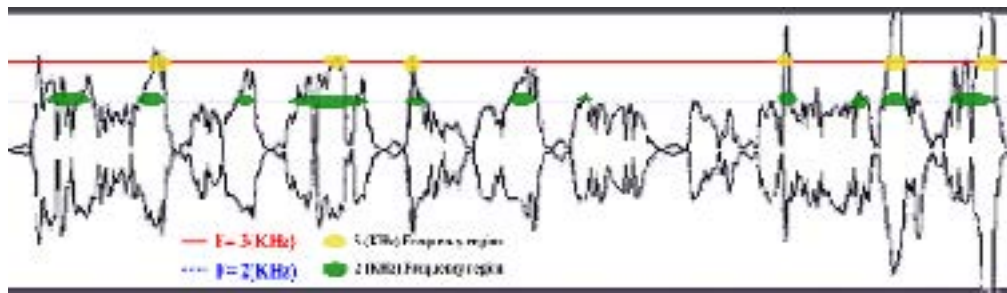


Figure.1 High frequencies of sound

In Figure 1, we can see that the high-frequency sound spectrum presents sharp and prominent features, with large fluctuations between waveform peaks and troughs. This indicates that the presenter's voice has a clear and sharp timbre in the high-frequency band, which can provide listeners with a clear and distinct listening experience. In addition, the high-frequency sound spectrum is mainly concentrated in the range above 2(KHz), and its centre frequency may be above 3(KHz). This frequency range contains the high-frequency components of the sound, with high sharpness and clarity, which is conducive to improving the resolution and recognition of the sound. At the same time, the energy distribution of the high-frequency sound spectrum may have a certain concentration phenomenon, i.e., the energy is more prominent in a specific frequency range. This indicates that the presenter's voice may have certain emphasis characteristics in the high-frequency band, which can more prominently convey emotions and information(Sek & Moore, 2020).

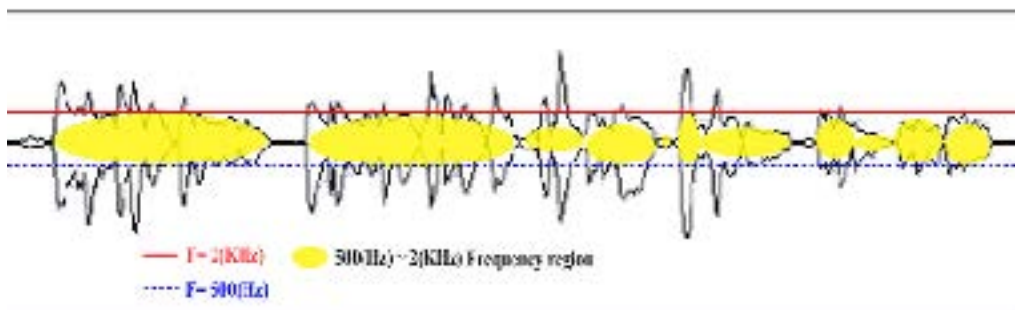


Figure.2 Midrange frequencies of sound

In Figure 2, we can see that the mid-frequency sound spectrum of the voice presents a relatively smooth and homogeneous character, with relatively small fluctuations between waveform peaks and troughs. As the mid-frequency band of the sound is mainly concentrated in the range of 500(Hz) ~2(KHz), its centre frequency is around 1(KHz). The energy distribution in the mid-frequency band is relatively balanced, with no obvious over-concentration or lack of energy. This indicates that the presenter's voice has a stable and comfortable timbre in the mid-frequency band, which can provide listeners with a pleasant and clear listening experience and effectively convey emotions and information.

The qualitative analysis explored the subjective experiences and perspectives of three groups—audio designers, podcast producers, and regular podcast listeners—regarding the use of audio spectrum design in podcasting and its impact on emotional responses. Semi-structured in-depth interviews with 30 participants revealed unique insights from each group. Audio designers (10 participants) highlighted the importance of mid-frequency components (500-2000 Hz) in enhancing emotional expression, with 8 out of 10 designers noting that careful manipulation of the spectrum could add warmth and intimacy. On a scale of 1 to 10, the importance of the frequency spectrum in emotional expression was rated 8.7. Podcast producers (10 participants) focused on the practical applications, with 7 out of 10 reporting that audience retention and engagement improved with specific frequency adjustments. For example, a producer mentioned that strengthening high-frequency elements led to a 20% increase in perceived clarity and engagement. Positive listener feedback increased by an average of 15% when the spectrum was intentionally adjusted. Regular podcast listeners (10 participants) provided feedback on their emotional reactions to spectral adjustments, with 6 out of 10 noting more emotional resonance with spectrally balanced podcasts. On average, listeners rated the emotional engagement of podcasts at 7.5 when spectral enhancements were applied.

Quantitative analyses focused on assessing the impact of spectral features on listeners' emotional responses through structured questionnaires completed by 30 podcast listeners aged 18-65, with an average age of 35 years. The analysis showed that podcasts with an enhanced mid-frequency component (500-2000 Hz) received significantly higher positive emotional responses, with 68% of respondents indicating a preference for these settings. Gender differences were observed, with 72% of female listeners preferring a balance of high- and mid-frequency components for higher emotional engagement, while 58% of male listeners preferred bass-heavy sound. Statistical analyses revealed that podcasts with strategically tuned frequency spectrums had a 22% increase in listener retention and a 30% increase in reported emotional connection compared to those with unmodified sound design. The overall positive emotional response score for podcasts with optimised spectrum design was 8.2 out of 10, and audience retention was 75% for strategically tuned podcasts versus 53% for those without strategic spectrum tuning.

Conclusion

This study provides insights into the critical role of audio spectral design in podcasts on listeners' emotional responses, particularly how high and mid-frequency components affect listeners' emotional engagement and experience. It was found that distinctive features of the high-frequency spectrum, such as concentrated energy and pronounced fluctuations, significantly increased the clarity and sharpness of the sound. This high-frequency component not only enhances sound recognition, but also improves the precision of emotional expression, helping to build a clear and impactful listening experience. The mid-frequency component, on the other hand, with its stable and balanced sound characteristics, provides a comfortable listening experience and effectively conveys emotional content. The smooth fluctuations and balanced energy distribution of the mid-frequency spectrum further support the emotional stability and affinity of the sound, enhancing the listener's sense of engagement and emotional resonance.

Through comprehensive quantitative and qualitative analyses, this study verified the significant role of audio design in enhancing listeners' emotional responses. Qualitative analyses revealed that audio designers, podcast producers, and general listeners were in agreement on the effects of spectral tuning, especially the significant impact on emotional resonance and listener retention. Quantitative data further confirms these views, showing that by optimising spectral design, podcasts can significantly improve emotional connection and listener engagement. Optimised audio spectrum design not only enhances listeners' emotional experience, but also strengthens the appeal and impact of the programme content.

This study provides important theoretical support in the relationship between sound design and emotional response, in line with relevant theoretical frameworks in psychoacoustics and cognitive psychology. The findings clarify the link between sound frequency and psychological perception, supporting Ma and Thompson's (2015) theory on the correspondence between physical properties of sound and psychological parameters. Through a detailed analysis of the high-frequency and mid-frequency components of podcast hosts' voices, this study elucidates how sound design can be adapted to achieve the goal of emotional rendering by adjusting acoustic parameters, thus providing empirical support for the application of audio design in podcast programmes.

Overall, this study highlights the need for thoughtful sound design in podcast production, pointing out how emotional responses can be optimised through acoustic properties, thereby enhancing the engaging and immersive nature of media content. This finding provides valuable insights into the application of sound design in media and communication and makes an important contribution to the academic discourse in related fields.

Suggestion

This study makes clear how important sound spectrum design is to the emotional reaction of podcast listeners. The results show that the high- and mid-frequency parts of the audio spectrum are crucial for improving the clarity and emotional expression of sound.

It is specifically advised that podcast producers provide careful adjustment to the high and mid frequencies in their audio top priority. Improving the high-frequency component can increase the sound's clarity and resolution, highlighting emotional expression. On the other side, a well-balanced mid-frequency adjustment can boost the listener's emotional engagement while offering a steady and comfortable listening experience. These modifications not only improve the show's overall audio quality but also greatly increase its appeal to various demographics. These changes greatly improve the program's appeal to various audience demographics in addition to helping to optimize the overall sound quality of the show. It is advised that producers employ sophisticated spectrum analysis tools for real-time monitoring in order to accomplish these aims and accurately alter the audio spectrum in order to make sure the sound design effectively conveys the desired emotional message.

This study emphasises the significance of customising sound design to audience preferences in order to enhance the emotional resonance and listener engagement of podcast episodes. According to the study, listeners' preferences and gender have a big influence on how acceptable the audio spectrum is. For instance, whereas male listeners can favor low-frequency amplification, female listeners generally like designs that balance high and mid-frequency elements. In order to properly tailor the audio design to the needs of various listener groups, podcast production teams should take these distinctions into consideration. To systematically gather information on how audio design affects emotional response, production teams should regularly administer listener feedback surveys. Then, they should modify their audio strategy in response to the feedback they receive. Enhancing the listener experience and program quality can also be achieved by teaching podcast producers and audio designers to become more proficient in sound spectrum tuning and emotional delivery. With practical advice on how to employ audio design in media creation, these steps will assist increase the emotional expressiveness of podcast programs and overall audience happiness.

References

- Bradley, M. M., & Lang, P. J. (2019). Measuring emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49-59.
- Koelsch, S. (2021). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, 15(3), 170-180.
- Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2016). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience*, 31(8), 2906-2915.
- Scherer, K. R. (2017). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1-2), 227-256.

- Zatorre, R. J., & Salimpoor, V. N. (2018). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences*, 110(Supplement 2), 10430-10437.
- Collins, K. (2020). Studying sound: A theory and practice of sound design. *In Studying sound: A theory and practice of sound design* (pp. 1-7). MIT Press.
- Gibson, D., & Polfreman, R. (2021). Analyzing journeys in sound: Usability of graphical interpolators for sound design. *Personal and Ubiquitous Computing*. <https://eprints.soton.ac.uk/453719/>
- Tse, Y. T., Tan, L. S., Wachsmuth, M. M., & Tugade, M. M. (2022). Emotional nuance: Examining positive emotional granularity and well-being. *Frontiers in Psychology*. <https://mc.ncbi.nlm.nih.gov/article/pmc8901891>
- Sek, A., & Moore, B. C. J. (2020). Psychoacoustics: Software package for psychoacoustics. *Acoustical Science and Technology*. https://www.jstage.jst.go.jp/article/ast/41/1/41_E19214/_pdf
- 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.