Volume 12 Issue 3 (2023)

Binaural Techniques, Functionality, and Applications: A Literature Review

Sophia Wang[1] and Julie Stoyka[#]

[1]University of Toronto Schools
[#]Advisor

ABSTRACT

Binaural techniques, the use of audio differences in the left and right audio channels, have been used with varying frequency over the past few decades. The main advantage of binaural techniques lies in how they create a sense of spaciousness and invoke specific audio scenes. This paper conducts a literature review about the studies surrounding recent binaural techniques and their functionality, as well as possible applications in the field of psychology. A dearth of research regarding recent binaural techniques was identified in the process of conducting the research for this literature review. Several opportunities for future research have been suggested, such as the use of binaural techniques across the vertical plane and evaluating the perception of sound and spatial scenes among the general public.

Introduction

The use of binaural techniques in music composition have remained a constant in the techniques composers and producers have used to invoke any sense of spaciousness in their music, even though the degree to which they have been used has been in fluctuation over the past few decades. In the context of this paper, binaural techniques refer to the use of different audio in the left and right channels of an audio recording. The introduction of electronic production and spatial audio editing has reintroduced this technique to more recent and modern composers and producers. Consequently, research and literature revolving around binaural techniques has also adapted and increased to reflect these changes. In light of this, this literature review aims to compile and summarize recent findings and studies of binaural techniques and their applications in several domains such as the reduction of stress and anxiety.

Upon conducting my research, I discovered that there weren't many published studies about the use of binaural techniques since the onset of COVID-19. I hypothesize that this may have to do with the fact that, before digital audio workstations (DAWs), all binaural techniques were composed of sound produced in different areas of a room, creating a spatial soundscape that could only be achieved in person. COVID-19 prevented the production of such music, and I believe that this may be the reason research on this front has halted. However, DAWs have since adapted to include many more binaural techniques. As such, I also believe that there will likely be an increase in research in the coming years (especially considering that artists have also moved to favour online platforms and electronic production, also as a result of COVID-19).

The general structure of this literature review follows my research process and goes along this general format: first, I explore the specifics of current binaural techniques and how music producers are currently using them, as well as investigate the current state of binaural techniques in modern music and their popularity among consumers. Following this, I conduct an investigation into the human perception of binaural music and spatial awareness. To finish, I look into the possible future applications of binaural music—in this case, in the field of psychology.

Research was conducted through search engines. Access to various studies were facilitated through the University of Toronto. In order to gradually expand the scope of my research, I started with the topic of binaural techniques in mind and, as I got a better sense of the basic techniques, began to research into the different applications of binaural music. My research focus was essentially shaped by the topics I encountered while conducting my literature search. Synthesis and inclusion into this literature review was done by categorizing the sources into the three categories that I laid out above: current situation, human perception, and application.

Section 1: Current State of Binaural Music

In order to first understand how binaural music operates and how people have used it in the past few years, I first investigated how binaural music is marketed to the public and how artists record binaural music in the industry. As such, I started with an article leaning more towards the professional than academic. "An Introduction To Binaural Recording" from *Sound on Sound* by Chris Korff is one of the more recent sources in this study, detailing music producers' current recording processes and how they incorporate binaural elements..

Korff believes that binaural techniques are becoming more and more relevant now in large part because of how more music consumers are using headphones as their main listening device. Compared to stereo speakers, binaural effects are much clearer through headphones, which is encouraging more musicians to use these techniques. Korff likens the human ears to omnidirectional microphones and explains how I know where sound comes from: if it reaches the left ear before the right, I know that it comes from the left, and vice versa. He also stated that my outer ears filter the sound in complex ways while it funnels into my ear canals; however, the complexities of this sound filtration are left unspecified.

This article mostly focused on how one might record binaural techniques, which were very valuable in understanding the models that came with later, more academic sources. Specifically, Korff introduced to me the Neumann KU 100, which was made specifically to stimulate the average adult human head, with two omnidirectional microphones in the places where the ear would be. In theory, whatever is recorded into these two omnidirectional microphones is what a person would hear if they were sitting there, allowing a listener to experience a much more accurate approximation of the in-person experience through recordings (provided that the audio recording is binaural and the consumer is wearing headphones or playing through two separate speakers).

However, Korff also mentioned that the Neumann KU 100 has its limitations, and proposed that a producer might decide to be their own "dummy head" instead, rather than relying on outside recording devices, though with the risk of an individual recording unwanted noises such as sneezing or coughing. Korff added that omnidirectional microphones do not "hear" the world the same way I do, but did not elaborate. I believe that this may have to do with the complex filtering he mentioned earlier in the outer ear and the ear canals, but this is only an inference. Further research corroborated Korff's statement that my ears filter and funnel noise, which I assume may not be present in omnidirectional microphones.

Building onto this, I decided to shift my focus to more studies and surveys conducted by researchers with the aim to investigate the current state of binaural music and techniques. "Binaural for Popular Music: A Case of Study" by Simone Fontana and his peers was one such source, and investigated the use of binaural techniques around its time. This paper was presented in a conference in 2007, making it less applicable to the current state of the art, but still provides valuable insight into the development of binaural techniques over time, as well as potential reasons for why binaural techniques' popularity has been in fluctuation over the past few decades.

Fontana raised the idea that the reason behind the unpopularity of binaural music relative to stereo music could simply be because people are not used to listening to binaural techniques yet, despite using headphones. Fontana hypothesized that this may still be due to the technological limitations of his time - many pop songs were still centered around stereo, which uses the same audio file for both channels, and most artists did not produce music with panned audio of any kind in 2007. Binaural effects were also rather difficult to gain while played on loudspeakers, and even though most of audio consumption is now through headphones, speakers still play a big role in creating ambience through music, in which spatial audio may very well be key in order to create the right atmosphere.

Fontana also raised the point that sound engineers may simply not be used to using binaural techniques, as stereo production has been the norm for a significant amount of time. According to him, binaural techniques at the time were very difficult to execute and required sound engineers to be receptive to these new techniques. Currently, from personal experience, it is much easier to use binaural techniques now than it was 20 years ago, especially in DAWs like Logic Pro. Panning is one such technique, which moves a section of the audio to the left or the right. Such techniques have gotten easier to use and more accessible, leaving Fontana's argument obsolete, but this is a good reference for what to expect in the future when advancing technology makes binaural techniques more accessible.

For other parts of the study, Fontana described the methods that he and his colleagues used in order to gather their own data. Five mixes in total were chosen for his study, all of which had different binaural characteristics. All participants were asked to evaluate three parameters, with which Fontana defined the spatial characteristics of the song: localization, sound relief, and spaciousness. The localization referred to how likely the listener was to identify where the sound came from. The sound relief was a measure of how likely the listener was to be able to reproduce the sound. The spaciousness was a subjective measure of how enveloped the listener felt.

According to Fontana, the participants of the study were often not very used to listening for "spacious" parameters—particularly, the spaciousness parameter gave even the musicologists many problems to identify. This corroborates the idea that binaural techniques weren't very familiar to people of that time. In order to get more people to benefit from the effects of binaural technology, Fontana suggested a gradual introduction of binaural music into the mainstream.

Other sources investigated the differences between headphone and speaker use. However, these studies often did not yield many results. One such study by Ian Eiderbo, named "How does binaural audio mixed for headphones translate to loudspeaker setups in terms of listener preferences?", aimed to investigate how binaural audio mixed for headphones translated to loudspeakers in terms of listener preference, and found that, overall, there was not a great difference in preference. The paper assumed that headphones were the most common playback system, and used this to justify the investigation into how compatible mixing for headphones would be when played on other systems.

Eiderbo also raised the concern of processing chains. Most binaural mixes require binaural processing in order to retain their original effects, which uses "crosstalk" between the left and right channels. However, he noted that many listeners don't use the software required to binaurally process media.

Eiderbo's paper was also the first source that introduced me to the Head Related Transfer Functions, or the HRTFs, which characterize how an ear receives sounds from certain points in space. He stated that the HRTFs are an important aspect of how we perceive and locate sounds, referencing the localization that I read about earlier in Fontana's study regarding how well participants were able to locate a sound source. However, while Eiderbo also mentioned the time difference between the arrival of sound at the two ears, he also mentioned amplitude difference, which was previously not mentioned in Fontana's study. Eiderbo stated that, since the human head naturally obstructs sound, the two ears perceive different obstructions of higher frequencies, thus providing a sense of directionality.

In the measurement of HRTF, Eiderbo stated that the specific HRTF is different for every person, as the ear and head shape varies from person to person. However, he also noted that there are approximate solutions for this, such as the Neumann KU 100 that I encountered in my first source. (As a side note, it appears that the Neumann KU 100 is the recording device of choice for researchers—I can see several more instances of the Neumann KU 100 being used experimentally in the following literature.)

Ultimately, the sample size of Eiderbo's research, while substantial, was composed only of students specializing in music technology. As such, the table of keywords that he compiled was full of technical terms that the general public is not very likely to understand, especially if they've never had exposure to modern music composition. This is one of the limitations of this study and its methods. Nonetheless, the study did find that there seemed to be no negative effect on the preference of the listeners when the mixed audio done for headphones was played through other playback systems.

Section 2: Spatial Audio Perception

Now that I had a relatively better idea of what the current state of binaural techniques looks like in popular music, I decided to look more into the psychology of a human's spatial audio perception, in order to better investigate the applications of binaural techniques. In this vein, I also looked at studies that compared a human's ability to perceive spatial audio with a machine's ability to do the same. One such study, "A Comparison of Human against Machine-Classification of Spatial Audio Scenes in Binaural Recordings of Music" published in *Deep Learning for Applications in Acoustics: Modeling, Synthesis, and Listening* in 2020 by Sławomir K. Zieliński and his colleagues, was very helpful in my review.

There were many more terms and definitions defined in Zieliński's paper that were extremely helpful. Specifically, Zieliński defined binaural beats—the name of the beat that forms as a result of hearing two different frequencies in two different ears, characterized as the difference in frequency between the two.

In all, Zieliński found that machines are typically much better at identifying specific spatial audio scenes that were invoked by binaural music. He stated that most of the research pertaining to machine learning in relation to binaural techniques has had much to do with the localization of individual audio sources, and then stated that this would ignore "higher-level descriptors," although he never specified what these higher-level descriptors would be. Zieliński also stated that his study focuses more on the "holistic approaches" to the characterization of spatial audio senses, which implied that he focused on how machines and humans localized the combination of sources, rather than each source individually. It was never explained why this would benefit the study, but it was a useful distinction to make.

Three spatial scenes were considered in Zieliński's study: the music ensemble located at the front of a room, at the back, and scattered around the listener in a horizontal plane. (I would like to ask if it would make any difference for a listener to listen to audio scattered around them in a vertical plane, although it was never considered or mentioned in this study, nor any subsequent studies. Various applications can be found with this, such as its use in big concert halls with tiered seating and box seats.)

In line with Fontana's study, Zieliński stated that spatial perception is greatly helped by interaural time difference, as well as interaural level difference (which was also a finding in Eidobo's study) and interaural coherence. Interaural coherence was never elaborated on, but I can infer that it refers to how well the different processed sounds in two ears match up with each other, perhaps to characterize the extent to which there is a level or time difference. (For example, less time coherence could mean that the sound source is likely to be directly on one side of the listener.)

Zieliński and his colleagues also brought up the cone-of-confusion effect, which occurs when a sound source produces identical sound delays and disparities for both ears. He stated that this effect could confuse a human's interpretation of these interaural cues, which is why humans exploit micro head movements.

The initial models Zieliński and his team investigated simply mimicked human binaural hearing and worked previously only under anechoic conditions, or, conditions with low reverberation of sound. However, he also stated that current algorithms for binaural processing have major drawbacks: algorithms often need prior knowledge about the characteristics of each individual source of noise, and many of these models rely on the previously mentioned micro head movements. While these are helpful, this study required the use of static head methodology. Zieliński went on to state that many studies currently include the classification of audio scenes—in contrast, his study aimed to have more organic settings rather than a tightly-controlled laboratory one. Any effect this might have had on the accuracy (with less-controlled variables) was unclear, but these scenarios were indeed closer to real-world examples and arguably with more practical applications.

The accuracy of humans was fairly low while classifying these organic audio settings, standing at around 42%. Zieliński and his colleagues hypothesized that this had to do with the static head methodology they adopted for this study. However, when the classification algorithms encountered scenarios that were never added into their training, they actually performed worse than the humans did.

As for corroboration in existing literature, Zieliński noticed that the automatic classification of spatial scenes based on binaural recordings by AI is still a fairly new field, and as such they did not have much corroboration in literature. The Convolutional Neural Network algorithm they used seemed to have similar accuracy to traditional algorithms. Zieliński stated that the algorithms' usability in unknown electro-acoustic situations (not just in anechoic conditions) needs to be improved, which could be a future path for study.

Other studies have also investigated how other aspects of binaural music affect the response of the human body and a human's processing of the audio recording. One such article, "Subjective diffuseness of music signals convolved with binaural impulse responses" published in the *Journal of Sound and Vibration* by Ryota Shimokura and his peers, investigated the spatial impression of audio. Shimokura split spatial impression into two subjective dimensions - the apparent source width (ASW), also known as the spaciousness of the audio recording, and the listener envelopment (LEV).

Shimokura also stated that during a musical performance, listeners do not listen to the impulse signal of music - they listen to the music itself. Moreover, the sound arriving at listeners is not only determined by the binaural impulse response (BIR) of the music hall in which the listener is listening to the music, but also the music signal. Neither of these two terms are further defined.

Shimokura's study investigated the time delay between the two ears together with the degree of constructive overlap of two audio signals and their relation to how much "subjective diffuseness" listeners report experiencing. They quantified these parameters with the Interaural Cross Correlation (IACC), which measures the difference between the signals received by the two ears of a person.

In this literature, music signals were characterized by five autocorrelation function (ACF) factors: the energy represented at the origin of sound delay, the delay time of the maximum peak, the amplitude of the first maximum peak, the width of the peak, and the degree of tonal components and the reverberation. However, their calculation of time delay was based on the measurement of the horizontal angle of each reflection of sound, which became very impractical. This was one of the limitations of this study.

Section 3: Applications in Psychology

In light of my findings, I then put my focus on how binaural techniques might be applied to other fields, such as medicine or psychology. Searching through academic papers, I found several articles that detail possible uses of binaural-based techniques in psychology, specifically in the reduction of stress. One such paper by Arianna Parodi and her colleagues, titled “A randomized controlled study examining a novel binaural beat technique for treatment of preoperative anxiety in a group of women undergoing elective caesarean section” and published in the *Journal of Psychosomatic Obstetrics & Gynecology*, focused on the role binaural techniques might play in improving mothers' positive experiences with childbirth in order to decrease maternal anxiety and possible postnatal complications.

Parodi noted that many studies have looked into ways to reduce stress and preoperative anxiety in adult patients, but few have focused on the binaural beat, which had been defined for me in Zieliński's study. It is worth noting, however, that Parodi defines the binaural beat as being a single tone with a frequency at the midpoint of the two different frequencies played in each ear, rather than the difference in frequencies.

In Parodi's study, a new binaural-beat-based technique was applied by adopting a novel algorithm developed by EfettoVIOLA Medical, called the Dynamic Multi Spectrum Phase Shift (DMSPS). The main benefit of such an algorithm was how it used infinite frequencies within bands in a dynamic manner. The developers of the DMSPS claimed that this would be able to generate the stress-reducing effect of the binaural beat phenomenon in just a few minutes, compared to the traditional binaural beat technique, which would need a treatment time of roughly an hour.

Patients were given either music, binaural music, or no music before their surgery. Parodi reported that there was not a statistically significant difference in pre-surgery anxiety levels between the different groups even after treatment, but there was a very noticeable difference in the difference between anxiety levels pre-surgery and post-surgery. Notably, binaural music made the patients significantly less stressed after the surgery compared to the control group, and even compared to the normal music group. This suggested that the DMSPS could give more relaxation to participants than the placebo effect.

Parodi's findings have the potential to improve the wellbeing of not only the many mothers who have to undergo caesarean section, but also any person who undergoes surgery. Nothing in the study suggests that these findings can only be applicable to mothers and their anxiety—in fact, the study characterizes preoperative surgery stress as the stress being investigated. As such, future studies should also consider verifying the findings of this study, so that hospitals may consider using binaural beat techniques in future stress-management initiatives.

Other studies have also investigated the binaural beat phenomenon in relation to relaxation. One such survey done by Siddharth Sharma and his colleagues, titled “Survey on Binaural Beats and Background Music For Increased Focus and Relaxation” and published in the *Institute of Electrical and Electronics Engineers* in 2017, investigated a proposed method of combining both short bursts and continuous binaural beats in order to help participants focus and relax. Unlike other studies, however, Sharma raised the idea that the music in question may have to be in a specific frequency range, depending on the brainwave that you want to strengthen. This would be important to the application of this technique, as different brainwaves encourage different behaviours.

Sharma noted one problem in the current state of the research into binaural beats for relaxation and increase of focus that had not yet been mentioned: systems that employ the use of only short bursts or a continuous stream suffer from predictability. Sharma believes that participants' anticipation of, or habituation to, the binaural beats prevents the desired effect from taking place.

As such, Sharma and his team proposed that, to counter this effect, they would combine both the short bursts and the continuous stream of binaural beats, making the audio less predictable and hopefully more effective. The binaural beats in the study were also superimposed with instrumental music that they defined as neutral in nature. The level of neutrality of the music can be questioned, as the emotions music can invoke are subject to one's own judgment; however, it is true that some music can be characterized as less potent in terms of emotion, even if it's not entirely neutral in nature. Nonetheless, such instrumental music is desirable when an individual wishes to focus and remain productive, as has been corroborated by several sources.

Sharma only suggested this proposed mechanism and has not yet tested it, which could be a direction for future study. Certainly, the proposed system has its benefits and merits some trials and experiments. Such a system would be relevant not only to students and workers who wish to remain productive, but also to music composers who are producing such music. Some producers aim to produce music that is calming and soothing, known as "lo-fi" in the popular music industry. Making use of these binaural beat techniques, if they're proven to be effective, can help producers gain the effect they wish—even more so considering that many lo-fi listeners listen to music through headphones, and I have established in earlier sources that headphones are one of the best playback systems for binaural music techniques.

Conclusion

In summary, the use of binaural techniques in music production, while significantly more mainstream than before, is still not familiar to many consumers, though there has been a recent increase in consumer products that take advantage of binaural audio, such as Apple's "Spatial Audio". Nonetheless, due to the nature of human sound perception, binaural techniques have many benefits and possible applications in psychology, especially regarding reducing stress and increasing focus.

In my literature review, I have identified some gaps in the current state of research. For one, Eiderbo's study of HRTFs could be done again with a sample size composed of the general population, in order to get a better understanding of the perception of sound and spatial scenes among the general public. There is also room for improvement in the characterization of music signals, which currently require an impractical amount of measurement of sound reflections (echoes). Newer methods would greatly benefit any future studies done on the spaciousness of audio and a signal's ACF factors. Additionally, there are many who might find it beneficial to test Sharma's proposed mechanism and evaluate the focus of an individual while listening to a mix of short bursts and a continuous stream of binaural beats.

As I have stated earlier, there is also the opportunity for research regarding spatial audio in the vertical plane. While not applicable to headphones, investigating differences in audio in the vertical plane could be applied to various settings, such as big stadiums or concert halls with tiered seats, or movie theaters with speakers placed at a higher elevation than the seats. Currently, research only focuses on the horizontal plane, which limits the options of sound designers who wish to make use of these binaural techniques in live music performances or playback.

The applications of binaural techniques to the enjoyment of music and the study of psychology are still relatively new fields, which makes binaural audio important for researchers to explore, as it may very well end up as a crucial part of music production in the coming years. While researching, I became aware of how little research there truly is about recent binaural techniques, with only 4 accessible studies I've found published in the past three years. Music production is not a field that researchers often study. However, when the body of literature grows, producers and sound designers can be better informed of what strategies work best, effectively enhancing the entire field of music production and allowing it to move forward in the coming years.

References

Eiderbo, I. (2021). How does binaural audio mixed for headphones translate to loudspeaker setups in terms of listener preferences? (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-85732

Fontana, S., Farina, A., & Grenier, Y. (2007). *Binaural for Popular Music: A Case of Study*. 13th International Conference on Auditory Display, Montréal, Canada, June 26-29, 2007. Retrieved from: https://www.academia.edu/47997100/Binaural_for_Popular_Music_A_Case_of_Study

Korff, C. (2023, June 1). *An Introduction To Binaural Recording*. Sound on Sound. https://www.soundonsound.com/techniques/introduction-binaural-recording

Parodi, A., Fodde, P., Pellecchia, T., Puntoni, M., Fracchia, E., & Mazzella, M. (2021). A randomized controlled study examining a novel binaural beat technique for treatment of preoperative anxiety in a group of women undergoing elective caesarean section. *Journal of psychosomatic obstetrics and gynaecology*, *42*(2), 147–151. https://doi.org/10.1080/0167482X.2020.1751607

Sharma, S., Rewadkar, S., Pawar, H., Deokar V., & Lomte, V. M. (2017). *Survey on binaural beats and background music for increased focus and relaxation.* International Conference on Emerging Trends & Innovation in ICT (ICEI), Pune, India, 2017. 10.1109/ETIICT.2017.7977018

Shimokura, R., Tronchin, L., Cocchi, A., & Soeta, Y. (2011). Subjective diffuseness of music signals convolved with binaural impulse responses. *Journal of Sound and Vibration, 330*(14), 3526-3537. https://doi.org/10.1016/j.jsv.2011.02.014

Zieliński, S. K., Lee, H., Antoniuk, P., & Dadan, O. (2020). A Comparison of Human against Machine-Classification of Spatial Audio Scenes in Binaural Recordings of Music. *Applied Sciences*, *10*(17), 5956. http://dx.doi.org/10.3390/app10175956