

Music and Sound for Media

Non-linearity and interactivity in game sound: surround & binaural

Immersive audio design techniques in general media, with focus on game development

Sounds evoke emotion which is key in maintaining the attention of the listener. If a game or film sounds boring or is full of silences, engagement is reduced and the mode of entertainment is less effective. With the global games industry generating a revenue of \$162.32 billion in 2020, there is a market for immersive entertainment. (Dobrilova, 2022). Engaging content means variety, and sound design is no different in that approach. Linear and non-linear sounds, along with different production techniques can be used for different effects. Non-linear sounds refer to sounds which are unpredictable in some form; often referred to as “the secret to a good horror or drama”. They are “too loud for the normal musical range” with examples including “the distress calls of wild animals, a child’s cry, and sudden and unpredictable frequency changes of acoustic instruments” (NPC, 2018). Studies published by UCLA reference the fact that “A variety of vertebrates produce nonlinear vocalizations when they are under duress” (Blumstein, 2010) which can evoke an increased emotional response from listeners; hence utilised within sound production.

In 2010 a study by Daniel T. Blumstein, a variety of films were analysed for use of non-linear sound with the results suggesting “that film-makers manipulate sounds to create nonlinear analogues in order to manipulate our emotional responses.” Horror films used more high pitch female screams and sad films “suppressed noisy sound effects” (Blumstein, 2012). Linking sound and visual media, the paper ‘The sound of arousal in music is context’ (Blumstein et al, 2013) concluded that “the presence of a video image suppressed the ability of simulated nonlinearities to modify arousal.” Meaning that without the video context to a harsher non-linear sound, the listener had a more intense reaction. Video therefore adds context to sounds and the two work simultaneously to create a real life image of what is being depicted within fiction. Research and experiments on the links between sound and visuals within gameplay have majority concluded that without sound, players lose a key aspect to the game, resulting in a helpless feeling combined with a reduced engagement rate. (Jorgensen, 2008).

“The term “diegetic” is used to refer to sounds that originate within the game world” (Robinson, 2019). An example of this is dialogue or footsteps. Diegetic sounds are subject to spacial design and effects such as reverb and echo. The sound should change depending on the space in which the character is contained within, with the example of a cave setting having both a high reverb and echo due to the sounds being reflected off the close surfaces. Using these effects creates a realistic setting as it’s sonically similar to a real cave setting, hence creates a realistic and engaging user experience.

Designing sound for spaces can also include non-diegetic sounds such as suspenseful music which is played in certain settings. Music for a cave setting is likely to sound more apprehensive and intense in comparison to a riverside setting. Effects such as distortion can create a sense of urgency and provide feedback to the character regarding an action, such as the more distortion on a punch, the harder the punch. Just like in music composition, sound design can also be used to create tension, with the likes of echo being used to define how big and empty the space is, hence creating tension in the unknown. Ambience within sound is also key in depicting the surroundings; ambient motions such as trees swaying or birds chirping can create a relaxing environment, whereas ambient environments such as strong winds or heavy rain can have the opposite effect. Reverb can create a sense of natural space, with the aim being to create a soundscape that reflects the real life environment. It can also be used to exaggerate certain actions, such as the thud of a door slamming to create a sense of dread. Camera angles and first person perspective can also play a huge part in the sound design, as the further the camera from the sounds, the quieter they will be portrayed in the game world. The use of non-diegetic sounds in video games can differ from linear media due to particular musical cues being used to notify the player, such as a sound cue for their declining health or music that symbolises a danger zone. (Volland-Butler, 2019).

Some sounds may be considered to be in the purgatory between diegetic and non-diegetic, hence transdiegetic. Transdiegetic sound usually uses dialogue to convey information to the player, such

as combat tactics in 'Warcraft III' where announcements aren't made directly into the game world but also not from outside the game. (Volland-Butler, 2019).

Looping and repetitive use of samples is heavily utilised within game sound design due to the inability to have infinite sound libraries for both storage and development reasons. Background music is looped, and within the likes of FMOD, loops can be altered slightly so as to create a new inversion of the loop, using a probability formula to switch between a selection of similar but different loops. Footsteps generally also have a small bank of sounds that are replayed 'in shuffle' throughout the game, often with slight pitch variations, as would occur in real life. A bank of footstep sounds can't just switch between 2 recordings as this would become repetitive, however also can't alternate between 100 different samples as this is unrealistic, particularly if the character is walking across the same surface. File space is also a consideration in the programming of sounds (Robinson, 2019) as graphics often have priority over sound and there is still a limit on how much file space games can take up. Technological advancements have increased space within sound development, allowing for a greater bank of sounds and music to be stored. In the past, games such as 'Gabriel Knight: Sins of the Fathers' (Sierra, 1993) used a single loop when entering a scene regardless of times entered. Today's technologies allow us to use a variation of loops to diversify the sound and create a less monotonous and more engaging game environment.

UX Sound Design and Feedback for Non-Linear Media

Non-linear media can be defined as media that can be interacted with by the user, such as a video game, website or social media. In UX design, sound design is key in aiding visual components. Sounds should provide both context and feedback for user actions, while also giving insight into threats or opportunities within the game world. Feedback is key in game immersion, as sounds are used to alert the player to various factors that will affect their gameplay: their environment, health, or proximity. The importance of interactivity in game play is directly linked to audio due to the nature of feedback, with interactions stemming from audio cues and the players need to respond. This response to the game is interactivity, and the more often this occurs, likely, the more the person is engaged and will therefore enjoy the game and keep playing.

An article published on UX Design (Foureaux, 2019) also highlights that the sounds used must align with the brand or game that it's being designed for, as if not, the sounds will feel out of place and result in a less realistic player experience. Sound placement is also important as sounds with harsher textures can be used to portray a more frantic message. Sound with more harmonic frequencies may sound eerie and can be used to evoke emotion, this is highlighted in the article "harmonically complex sounds indicate priority", meanwhile sounds that provide feedback that aren't high priority within gameplay shouldn't use complex sounds as they are more attention inducing, therefore are not appropriate for low priority actions. In terms of usability, key principles are often referred to as heuristics rather than set rules, with industry practitioners such as Jakob Nielsen creating the "10 Usability Heuristics Principles" (Nielsen, 1994, pp153), to be used as a guide to UX sound design.

Design concepts such as skeuomorphism are also referenced, whereby virtual sounds mimic real life action sounds, such as a computer's recycling bin making a crumpling paper sound or online notes looking like lined paper. (Foureaux, 2019). Skeuomorphic design is often why digital plugins are designed to look like older analog counterparts. This can be incorporated to ensure realistic impressions for the player, with the idea of skeuomorphism gaining popularity with the rise of VR.

Surround & Binaural

Sounds can be heard in a multitude of formats. From mono meaning one audio stream/sound is duplicated into each speaker; stereo, meaning there are two channels of audio that can differ such as being panned, then there's binaural and 3D sound. Binaural sound refers to the way the average person can hear, using 2 ears to create a sense of direction. Sounds can come from above, behind and from each side. This is different from stereo as only sounds can come from either left or right. 3D sound is the most recent advancement in terms of sound, with it becoming

popular due to virtual reality. When a character moves around in a virtual world, they want the sounds to move with them from any angle.

“Binaural recordings, however, are optimised for headphones and can recreate the perception of distance - making your audio more immersive” (Smith, 2017). The idea of binaural recording has been around for a long time, with Neumann releasing the KU-80 in 1970. This had 2 microphones placed on either side of a dummy head to stimulate real human hearing, however for musical recordings, it never took off due to a few limitations and was widely regarded as a failed technology within the decade. Problems with recording capabilities meant that the microphone wasn't diverse enough to record a variety of instruments, as often a different microphone setup would be used for drums compared to the likes of vocals. (Danso, 2017). The book 'Music, cognition, and computerized sound: an introduction to psychoacoustics.' (Cook, 1999, pg90-101) talks about how with binaural sound, often the user can't distinguish if something is in front of them without visual stimulus. This book was published long before realistic VR became popular or possible, hence with today's technology, binaural sound can create a realistic balance with video elements to fully immerse the listener/player. This is reinforced through the Helmut Haas Effect, a psycho-acoustical effect whereby the first sound wave to reach the human ear is usually used to identify distance, while the second, even if received with a slight time difference, is perceived to be the same sound even if it's louder and arrived with a minuscule time difference.

It's widely debated whether binaural sound is worth being implemented over surround sound in virtual or video environments. Surround sound still enables a 360 sound field around the listener, with sounds being coordinated from onscreen events to the relevant speaker, such as a gunshot on the left being heard louder from the left speaker. Dolby Atmos, one of the most recent iterations within sound technology, also utilises sound systems from above, similar to binaural sound, creating an even wider field of sound, however, academics often argue that “recreating exactly what your ear hears in any given scenario, binaural audio accomplishes a far more natural 3D sound without a costly armada of speakers.” (Lalwani, 2015). Binaural sound is however much easier to implement through headphone applications, hence is more useable paired with VR rather than in film and TV.

3D sound application within VR enables another dimension of realism as it can overcome one of the major design challenges in relation to sound. VR users have the capability to look in any direction and may not be facing the most beneficial way when a key event occurs, hence when 3D audio cues are used it can quickly veer the users attention to the event through sound.

Musical applications of binaural sound include performances from the likes of Coldplay, Imogen Heap and Thom Yorke from Radiohead who was pictured with a binaural setup in 2007, however the end result merely sounded like a more spaced out version of a stereo recording. Chris Milk, a video director, has produced many 360, immersive concert experiences with the most prolific being "Hello, Again," a revolutionary VR film of a Beck concert. Users could look in every direction and the sound system maintained the capability to produce an accurate representation of real life dynamic sound. For the creation of this concert experience Milk “devised a system that used an omni-binaural head with eight ears, with four pairs of 3Dio's Free Space microphones implanted at every 90 degrees.” (Lalwani, 2015). Algorithmic applications also have binaural capabilities such as 'Newendo', a sound editing software used in post production to create a similar effect to binaural recording. (Smith, 2017).

Within games, 3D sound/Binaural sound can create context for players due to the real world sound alignment. Games that have the option for binaural sound include 'Sniper Elite 4', 'Assassin's Creed: Origins' and 'Resident Evil Village'. Graphics have often been favoured over sound, hence binaural or 3D sound specifications are often ignored by game players with the focus on what they can see rather than what they can hear. This may be the reason that sound technology such as binaural audio have been around for a long time without implementation in the gaming world. The problem for audio designers with integrating various playback options for users is that there needs to be the option to switch to meet the needs of each user, as some may have surround sound speakers while others may want a game that has a binaural capacity; however, when one system is used with the wrong setup it can reduce engagement and the intended effect.

New technologies in game sound are generally focused on 3D and binaural audio. With game designers implementing this into their games, there's been an increased demand for headsets that can match these capabilities. Sony has recently released their PS5 Pulse 3D wireless headset which has "Tempest 3D AudioTech", meaning their headsets are now capable of binaural audio, claiming to be able to hear rockets zooming past one ear and "where it feels as if the sound comes from every direction". (Playstation, 2022). Along with binaural audio capacity, this new headset has a built-in 'rumble motor' "to provide additional feedback during gameplay. This means you'll be able to physically feel the impact of an object flying past your face. Research also suggests that including this type of force feedback can help with motion sickness" (Playstation, 2022). This gives sense of VR without actually using a VR headset due to being capable of feeling the effects of the game.

Within gaming, most diegetic sounds should or could be binaural, such as players movement and distant sounds. Environmental settings are also effective in binaural audio, however may be less important, for example, an enemy approaching from your left would be of higher priority than birds chirping and the wind blowing from the right. Narrative or story telling aspects of games tend to use ordinary stereo audio unless there is physical representation of a story telling character within the game.

Although sound paired with visual media generally creates an immersive environment, there are occasions where sound can hinder the intended affect and distance users from the experience. In the paper 'Transdiegetic Sound and auditory immersion in an asymmetrical cooperative game' published by the University of Lincoln (Volland-Butler, 2019) it refers to how "in a multiplayer context [...] communicating with other players can break in-game immersion". Theories based on this research suggest this might be due to human vocalisation not fitting in with the virtual world due to its associations with reality, with others suggesting that interacting with other players takes attention away from the game, hence the player may be less focused on the game and even less on the sound and sound quality during communications.

Linear vs Non-Linear Media

A report done in collaboration with Meta in 2016, concluded that VR and binaural audio allow for sound to be 100% acoustically accurate; however, for non-linear media such as film, binaural audio has the capability to be 100% controlled. (Bible, 2016). Binaural audio outside of VR generally means the sound is not adjusted when the listener moves their head, but maintaining a 360 soundscape while looking at the screen. Since screens are stationary, this maintains effectiveness.

Binaural sound creation in films with its increased control over VR sound design, allows for specific aspects to be tailored to the listener in each scene, rather than in VR, where these factors are automated during production. Thomas Bible's article 'Binaural Audio for Narrative VR', refers to the audio production process for the short animations Henry and Dear Angelica. Mixing techniques used within binaural films differ slightly from games due to this control; "too much sound can be overwhelming to a listener", including spatial sound. In Henry, we addressed this in busy moments by collapsing sounds that were far away from the listener (but close to each other) down to single sources." (Bible, 2016). Various attenuation lines were used in relation to volume and distance from the character, with some being cut during busier scenes. This would be harder to implement in games as it would mean planning for unlimited circumstances due to the nature of non-linear media. With these particular short animation examples, music was left as 2D audio "since most people are used to hearing music on headphones while also hearing sounds from the outside world." However, an issue arose when the binaural sounds around the listener changed in volume, the music was no longer balanced in the mix. (Bible, 2016). In a different example, one scene contained binaural music, with stems being used in different corners of the room, however the key aspect to this example is that the room was black surrounding the character, meaning the sound didn't appear to attach to any object, having so-called 'floating phantom sources'. When using this technique in a visible room "the phantom source issue became much more noticeable and jarring." (Bible, 2016). This links back to the Helmut Hass effect whereby people find it hard to gauge where an object is without forward visual stimulus, however in a black room, there is no location to gauge so the sound is simply all around.

Conclusion

To conclude, immersive audio techniques, particularly within non-linear media forms are central to developing the experience. Without sound, visual stimulus would be less engaging, hence appropriate sound design is crucial in game design. I would argue sound is even more important within non-linear media forms such as gaming due to the feedback aspect. Feedback within sound design gives the player knowledge and context, both of which are needed to play the game. Although an immersive sound experience can be achieved through a variety of industry standard techniques and stereo audio, the use of a surround sound or binaural setups can greatly improve the overall experience. In my opinion, binaural audio is the most effective sound form for engagement in non-linear media due to its realistic depiction of audio, bringing a real world aspect into the virtual world, hence its popularity within VR. In comparison, within linear media, binaural audio may not be the most effective due to the movements of the user remaining unreflected. Dolby Atmos and new sound systems also provide an immersive sound experience, however, I believe if binaural audio is done correctly within game development, using modern audio production techniques, it can be the future of sound for gaming.

Bibliography

Nevada Production Crew (NPC), 2018. 'PRODUCTION NOTES: NON-LINEAR SOUNDS' (Accessed: January 2022) [Available at: <https://nevadafilm.com/production-notes-non-linear-sounds/>]

Daniel Blumstein, 2010 'Do film soundtracks contain nonlinear analogues to influence emotion?' Royal Society: California. [Available at: <https://royalsocietypublishing.org/doi/10.1098/rsbl.2010.0333>] (Accessed at: January 2022).

Daniel Blumstein, 2012. 'The sound of fear' TEDxUCLA [Available at: https://tedx.ucla.edu/talks/dan_blumstein_the_sound_of_fear/] (Accessed: January 2021)

Daniel Blumstein, Gregory Bryant, Peter Kaye. 2013. 'The sound of arousal in music is context' (Accessed: January 2022) University of California: Los Angeles

Ciaran Robinson, 2019. 'Game Audio with FMOD and Unity' 1st Edition, Routledge.

Smith, 2017. 'Binaural audio: What is it? How can you get it?' (Accessed: January 2022) [Available at: <https://www.whathifi.com/advice/binaural-audio-what-it-how-can-you-get-it/>]

Danso, 2017. 'The binaural failure of the recording industry' (Accessed: January 2022) [Available at: https://medium.com/@andrewdanso_84496/the-binaural-failure-of-the-recording-industry-b60bf61c6781]

Lalwani, 2015. 'Surrounded by sound: how 3D audio hacks your brain' (Accessed: January 2022) [Available at: <https://www.theverge.com/2015/2/12/8021733/3d-audio-3dio-binaural-immersive-vr-sound-times-square-new-york>]

Cook, P. R, 1999. 'Music, cognition, and computerized sound: an introduction to psychoacoustics.' Cambridge, MA: MIT Press, pg 90–101

Foureaux, 2019. 'Sound Advice: A Quick Guide to Designing UX Sounds' [Available at: <https://www.toptal.com/designers/ux/ux-sounds-guide>] (Accessed: January 2022).

Nielsen, J, 1994. *Enhancing the explanatory power of usability heuristics*. 1st ed. Boston: Human Factors in Computing Systems. [Available at: https://static.aminer.org/pdf/PDF/000/089/679/enhancing_the_explanatory_power_of_usability_heuristics.pdf]

Barker, 2021. 'PSVR2 Guide: Everything You Need to Know About PS5 Virtual Reality' [Available at: <https://www.pushsquare.com/guides/psvr2-guide-everything-you-need-to-know-about-ps5-virtual-reality>] (Accessed: January 2022).

Sony, 2022. 'PULSE 3D™ Wireless Headset' [Available at: https://www.playstation.com/en-gb/accessories/pulse-3d-wireless-headset/?emcid=pa-co-422388&gclid=Cj0KCQiAuP-OBhDqARIsAD4XHpd3uz74OZq_WiHA8EaZScEe68mzTKMDvk66j73DHUztLNQSuDtFcr8aAlfbEALw_wcB&gclidsrc=aw.ds] (Accessed: January 2022).

Bible, 2016. 'Meta - Binaural Audio for Narrative VR' [Available at: https://www.oculus.com/story-studio/blog/binaural-audio-for-narrative-vr/?locale=en_GB] (Accessed: January 2022)

Volland-Butler, C. 2019. 'Transdiegetic Sound and auditory immersion in an asymmetrical cooperative game'. *University of Lincoln*, Issue 1, pp1-15. [Available at: <http://eprints.lincoln.ac.uk/id/eprint/39311/1/VollandButlerCharlieComputerScienceOctober2019.pdf>](Accessed: January 2022).

Jørgensen, K. 2008. 'Left in the dark: playing computer games with the sound turned off', Farnham: Ashgate. pp. 163–176.

Dobrilova, 2022. 'How Much Is the Gaming Industry Worth in 2021? [+25 Powerful Stats]' (Accessed: January 2022) [Available at: <https://techjury.net/blog/gaming-industry-worth/#gref>]