This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2996371, IEEE Computer Graphics and Applications

There's more to Sound than Meets the Ear

Sound in Interactive Environments

Review Submission

Abstract — How important is sound in an interactive environment? For example, what happens when we play a video game without sound? Does the game still have the same impact? Even if sight is the primary sense in interactive environments, sound is also important, and should not be overlooked during the development process. The necessity of sound for perceptive quality enrichment in virtual environments cannot be underestimated. However, how designers should integrate and leverage the benefits of sound design effectively in an interactive environment can be challenging. This short article, discusses a variety of important and intriguing psychological concepts and immersive sound techniques, used in interactive environments, such as video games, to improve engagement and enhance the experience (from passive background music to active and procedural sounds). Computer graphics has proven itself in many fields of entertainment and computing as a means for communicating and engaging users (visually). This article discusses the hidden abilities of sound in interactive environments (e.g., the emotional, subconscious and subliminal impact). We explain how different sounds can be combined with visual information to help improve interactive conditions and stimulate the imagination, not to mention, control (or steer) the user's emotions and attention.

Index Terms—Sound, Interactive, Unconscious, Design, Immersion, Procedural, Experience, Engaging, Video Games, Emotions, Attention

*

1 Introduction

Experience Sound is the second most important human sense after sight (Figure 1). Needless to say, sound offers a great deal of power (in the right hands) for controlling experience. Yet in interactive environments, such as video games, sound is underutilized in certain aspects. While sound and entertainment has always gone hand-in-hand from the earliest times, we are still making new discoveries about sound and is an active area of research in many fields, including video games [1]. Of course, sound in video games is still a pivotal factor of feedback and engagement (e.g., music and noises are able to trigger reactions, memories even emotions). The fact that sound help players with engagement, of whatever purpose and value is clear. This notion has taken hold in today's world with real applications (even outside video games in a variety of other contexts, such as, training-simulations, educational-games and rehabilitation programs).

We frequently use sound to convey all sorts of information and to help this information be remembered. This article discusses the role of sound in and around video games and its impact on engagement, mood and cognitive performance. Of course, sound and audio in games is a broad topic and would be beyond the scope of this article, however, we focus on a few interesting and insightful areas that have and are making an impact in recent years. Specifically, in the areas of sound design that are underutilized, such as, the psychological context between audio-visual senses, frequencies and emotional impact [2]. We review past and future trends in sound design, including recent advancements, like machine learning and procedural generation of audio content, which will have a substantial impact on the next generation of video games.

What is it about 'Sound' that is so Important? Sounds are all around us. Sounds provide information about the world we occupy. Sounds provide a description of the space we are in and describes most of the other things that occur in that space (for instance the noise of a clock ticking, a car engine rumbling or even the echo of a library hall). Hence, if video games want to successfully captivate and immerse players, they need to include sounds. Video games are not only a visual experience. They are also interactive (compared to other entertainment mediums, like television or the cinema) and work best when we take advantage of multi-sensory input, like auditory and touch senses. This is important, as multi-sensory information, such as sound, tells the player what is happening outside their field of view, and what location those events are taking place (in addition to any other background details related or not related to the events). Sound adds additional sensory information that confirms the events, and the actions that a player is performing in the virtual world, making the video game seem more plausible and realistic. A lack of sound (or poor sound) has a big impact on the player's quality of experience (in a video game) both consciously and sub-consciously.

Frequencies and the Affects on Emotions and Behaviours

Sound has the power to influence our emotions and behaviours [1]. A sound contains a multitude of information embedded within frequencies and signal patterns. Investigations and studies have explored the influence these underlying frequencies have on the mind [3]. Of course, frequencies alone do not impact our mood and behaviour, they must also be combined with other factors (e.g., volume and context). Speech is one example, how we speak has a big impact on what we mean and how it will be interpreted. The mixture of high and low frequencies with game play can be a powerful resource in reorganizing the player's mood. As we combine low and high frequencies together through music, a user's behavior can be steered towards a specific state (such as calming, focused or excited mindset). The challenge is finding a suitable 'mix' of high and low frequencies (in a specific pattern and/or context), as different types of sound might work for one game but not another due to the genre of the game. An example of a good balance is a military march. They play low frequency sounds (the drums) to make the soldiers want to move, yet can have a tiring influence over time, which is counteracted by mixing in high frequency sounds (the trumpets) to stimulate the brain and provide energy so the soldiers can march longer.

Contribution The key contributions of this article include (1) selective review and discussion on various audio design aspects in the context of interactive environments and their associated techniques, limitations and applications; (2) we also identify a number of important gaps and areas for future exploration that are often overlooked in sound and interactive environments; and (3) details on cross-disciplinary sound discoveries which would have an impact on immersive sound design (emotional, subconscious and subliminal sounds).

2 Cutting through the Noise

Why do our ears pick up when they hear something interesting? In a crowded room a specific word, sound or song can instantly catch our attention or change our mood. Similarly, while listing to the radio in the car, our mind will wonder, yet be **subconsciously aware (listening in the background)**, letting the music, discussions and news entertain us through a delicious recipe of intrigue and unknown. As broadcasters and producers invest a huge amount of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2996371, IEEE Computer Graphics and Applications

time into understanding the power of sound, understanding small changes in volume or how the order of sounds played will influence, tone, attitude, focus and more [4]. Even in video games, background music and environmental sounds add atmosphere - this atmosphere (even if subconscious) influences how we feel, react and respond. One of the most common sound's we will hear in a virtual scene (e.g., room or corridor) is the background sounds, such as, humming of lights, radiator pipes or mumbled talking. These subtle noises create a feeling (aura or mood), while the player intently listens for noises. Of course, these sounds need to be complemented with visual aids to be truly effective (since it would be strange to hear the sound of an elephant while walking down the corridor in a space ship - triggering confusion and breaking any immersive connection between the game and the player). What we must remember, is background noises works **synergistically** with other senses.

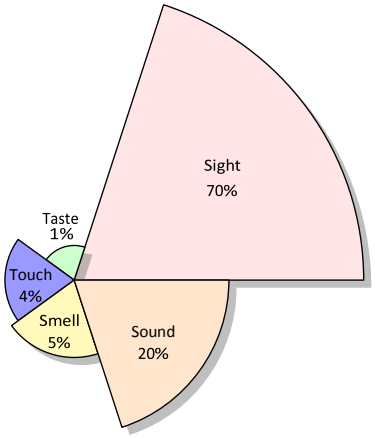

Fig. 1: Senses - Comparison of the weighting importance of the five human senses, sight first and sound second [24] (relative significance of senses to object perception).

Stimulating Sounds `Non-linear` sounds are an important and unique category of sound with a powerful means to stimulate or shock the user (for example, a weapon shooting in a video game). These sounds are classed as non-linear as they exceed the normal musical range of an instrument (or the vocal cords of living creatures) [5]. They can also be produced by sudden frequency changes in acoustic instruments. In the real-world, we often encounter them in animal distress calls (which may be the reason they produce a seemingly-ingrained reaction when we hear them – hard wired into our instincts).

In video games, these non-linear sounds are common way to inform the player that something is happening (often that they are in danger or they need to prepare for action) [6]. Obviously, this can also be complemented by altering the background music – such as, adaptively altering the speed or tempo to inject a sense of feeling into the game-play and experience (peaceful exploration vs high-speed action). These musical shifts add to the immersiveness of the overall experience, ensuring the sound always corresponds with what is happening visually on-screen. Recently, designers have been taken with the idea of using adaptive music to generate soundscapes dynamically - shifting soundtracks that react to the in-game action, creating new and unique music every time they are played [7] offering an endless amalgamation of audio-visual experiences.

Passive vs Active Sounds A passive sound is something we hear in the background, such as, music or the wind and may influence our concentration or mood. While an active sound engages our conscious awareness, for instance when someone answers a question correctly or incorrectly during a quiz after the 'buzzer' noise. Certain sound qualities are indicative of change, such as sudden or extreme sounds, sounds explain the stimulating and relaxing effects of music, and how mere sounds may induce changes in our thinking or emotions [8]. While there has been a strong debate on the impact of active sounds compared to passive sounds on mental awareness, there is no real scientific evidence. However, Gemba and Sasaki [9] showed some remarkable insights on how a brain is influenced and processes active and passive sounds. The work on active audio was done with 'monkeys' and demonstrated a strong underlying pattern of cortical activation (i.e., stimulation of the cerebral and/or cerebellar cortexes through attainment of sensory arousal during mental tasks). This evidence means that video games that use a large number of active sounds during the game-play trigger more sensory arousal that stimulates the experience (compared to primarily passive sounds).

Impact of Sound There is a growing body of empirical data in the literature [10,11,12], which focuses on the **influence sound has on our tastes, expectations and our experiences, not to mention our enjoyment**. For instance, one researcher, Charles Spence [10] compares sound to one of the lesser forgotten senses and has collected over 70 years of empirical data relating to sound and senses (taste and experience). This work was also supported by Bolmont [11] who in a related study demonstrated the ability of **sound to balance and control mood states and anxiety** in human subjects. Expanded investigations have also explored the combined impact of audio and visual stimuli, for instance, Tremblay [12] and other academics, have pointed out that the simultaneous presentation of in-congruent audio and visual stimuli can lead to **illusory precepts**. However, the common sense of belief surrounding the effectiveness of multiple modalities in immersive environments (audio and visual) is still an active topic of research. Of course, considering everything, we must not forget that sound is highly dependent on experience and is a controversial subject, especially in the area of experience and emotional impact.

Ways of Learning through Sound Sound, games and interactive environments go hand-in-hand with learning. For instance, Cakir [13] successfully combined the use of video with audio-visual material for foreign language teaching classrooms. Other researchers have applied audio with short video clips based upon popular cultures to enhance the classes' involvement and interest. This showed the challenges and problems of designing and using sounds effectively in communicating concepts. Stemming back to the fact that similar sounds would be perceived differently by different individuals resulted in contradictory qualities when used in immersive learning environments. Nevertheless, a useful aspect of sound with broader applications with less dependency on experience is the ability to condition the mind (to subconsciously calm or help individuals focus). As mentioned earlier in the article, certain auditory patterns and sounds have an instinctual impact (soft slow noises verses high-pitched shrieks). When these sounds are complemented with other material (auditory dialog or feedback sounds), they offer an improved learning process (and experience).

Sound Recording and Playback Video game sounds (even film sounds) are required to imitate the real-world. Often, large databases of audio recordings are created to capture real-world sounds for play back in virtual environments (from car engines to the howling wind on a cold winters night). The majority of the time, when they are played back in virtual environments, we are able to identify with these sounds and come across as `realistic'. Yet there are times, when these playback sounds do not capture the essence of the real world experience. Often due to our ears ability (and experience) in identifying `gaps' or `strange' anomalies with the playback sound. In the early years and even today, there are subtle differences between what is recorded and what we hear when we playback sounds. For instance, most action sounds do not actually sound like what the listener would have expected the sound to have sounded like. A good example are footsteps. If we try to record a real person walking, it would sound 'false' (or under emphasized)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2996371, IEEE Computer Graphics and Applications

to the listener. So to resolve this, the sound is recorded on different surfaces (even custom shoes) or digitally enhanced (sharpened or reverberated) to make the sound fit the context of the video games scene. This even leads onto an interesting area of sound akin to a topic of research related to realism and perception in graphics known as the `uncanny valley' [25]. Specifically the `audio uncanny valley', which was first suggested by Francis Ramsey to articulate responses to simulated sound, and the perceived "naturalness" of said sounds. As the audio uncanny valley defines how certain sounds affect a persons' perception, unpleasantness and negative affects, which influence the individual's emotion, and audio-visual synchronization (like when the lips and voice do not match).

3 Sound Design

Good and Bad Sound Design Often, audio is an afterthought in many interactive environment projects, with the assumption that sound is not the reason the project fails. Sound (like all aspects) should be designed, whether intentionally or not. In most cases, **bad sound design is worse than no sound design at all**. Good sound design is magic, while poor sound design is upsetting. When good sound design works, it brings the experience alive, works in synergy with all other aspects underneath. Often times though people are rushing to meet deadlines and have sound thrown in at the end. Examples of how sound design fails:

- Positional audio appears non-existent (important in stealth games), i.e., difficult to tell where voices or footsteps are coming from. Actually hearing them is not the issue (unrealistically loud) but always sounds like the sound is coming from right in front of the player no matter where the sound is supposed to be coming from
- Volume mixing is inconsistent and random. Actions and weapons have sound soft and voice audio is too low (forces player to turn sound off or turn the volume up far higher than usual)
- Actions (collisions and hits) and weapons sound weak in comparison to everything else. Some sounds are barely audible while subsequent sounds may be too loud - especially, if the player turns the volume up to compensate for lesser sounds)
- Actions with no corresponding audio (e.g., the player is hit or injured (shot in the back) but no audio indication
- Sounds which contradict one another in tone and style of the environment/scene
- Voices going through multiple walls and popping in/out when the camera moves around (sounds cutting out instantly instead of fading away when the player navigate the environment or leaves the rooms)

Can We Trust Our Ears? (Audio Illusions) Audio information has an impact on our 'mind'. Sound is able to alter our perceived vision and vice versa. When sound and visuals work together to alter our perception, this effect is known as the McGurk effect [14]. The McGurk effect is a cross-modal illusion usually resulting from conflicting information coming from different senses, namely sight and hearing. Alternatively one audio specific illusion worth noting that creates the illusion of infinitely increasing tones, is the Tritone paradox [15]. The Tritone paradox is achieved when a pair of Shepard tones are played sequentially but separated by an interval of a triton (or half an octave). This gives the listener the perceived illusion that the sound is continuously increasing (or ascending). This audio experience gives the listener the feeling of an endlessness (similar to travelling down a tunnel without an end) – a paradoxical sound which seems infinite while appearing to constantly change in pitch, though it actually cycles continually. These audio and visual illusions (tricks) allow us to control the users' senses to create engaging experiences that tantalise their imagination (creatively seeding gaps in the users mind and ways of thinking).

Sound of Silence in Video Games How do noises (like background ambience) impact a video game? A player can either be positively or negatively affected depending on how they perceive the noise. For instance, the background noise could be music, it could be fast beat and uplifting, encouraging and motivating the player to try harder (giving focus). Yet, the noise of crowds booing combined with slow rhythmic acoustics, could bring up negative feelings, clouding their judgments and physical responses. Other times background sounds could be distracting and annoying (so their needs to be a careful balance).

Sound in interactive environments offer a powerful engagement tool. However, we argue that sound should be used 'carefully' and as an aid (i.e., as a garnish used to embellishment accompanying activities and information - we do not want to add sound just because we can). Work by Morgan [16] on the effects of sound as a distraction upon memory inspired other investigations and studies on sound, including the impact of both speech and nonspeech intersensory perception on individuals [12]. Researchers, such as, Kalyuga [17] argued that sound with a text or picture is **not beneficial for learning**, by describing two cases where combining audio explanations with visual instructions had negative rather than positive or neutral effects. These results were explained as a consequence of working memory overload with guiding principles in the design of multimedia instruction are suggested. Really depends on context, as Zembylas and Michaelides [18] emphasis on traditions of silence while exploring the pedagogical value of silence in education, learning and reflection. The **context on when and when not to use sound** in learning is an open area, for example, if students are reading material from slides or listening to a talk, then other complementary sounds would probably be a disturbance. Similarly, in a game, if a player has to read information, distracting sounds make this more difficult.

4 PROCEDURAL

How Sound has Evolved Video games and interactive environments have evolved over the past few decades (see Figure 2). They go beyond repetitive coin collecting genre, towards the ability to embrace and tell complex stories, in a more cinematic aspect. As the design and graphics have changed, so has the incorporation and use of sound. Game designers use sound and music to make games more immersive and give the player' a more emotionally-charged experience. Not only is sound changing in use, but also how we store and generate it. We are moving away from pre-recorded libraries to more procedural approaches to address various limitations and offer new opportunities (in terms of realism and flexibility).

Why Procedural? Content generation constraints, such as, costs, time and hardware limitations (physical memory and network transmission times). Increasingly, consumers want higher quality content for their immersive environments, with more complex, realistic and detailed sounds, this is also compounded by shorter production times (which makes `manual' solutions unviable and encouraging `smarter' more automated and efficient approaches to be developed). In addition, there is also the factor of variations and modifications, having to update large libraries each time a small change is needed (such as in dynamic environments, these changes could be dependent on gameplay or situation) [20] and could be problematic if user intervention is required.

Could all Sounds be Done Procedurally When we say procedural sound/audio, we mean the audio content is generated by a set of computational functions or algorithms. This depends what sounds we want to create procedurally, and if they need to be totally procedural or use other ad-hoc methods (blending or randomization). For instance, audio sounds that are repetitive (like footstep and impacts) are ideal candidates. There are also sounds

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/MCG.2020.2996371, IEEE Computer Graphics and Applications

which have simple audio patterns (wind and ocean waves). There are more complex sounds which have a large number of variables/control aspects, like creature vocalizations and speech. Procedural approaches for generating sound has changed tremendously over the past few decades. Traditional approaches for generating procedural sounds focused on repetition and mathematical models, such as, with Fractals and Perlin noise [20]. These produced attractive solutions (for simple situations). As time has progressed, computational processing power and new techniques have allowed us to explore other avenues, such as, data extraction and training models (machine learning). These recent approaches have allowed us to develop procedural sounds that embrace real-world examples (learn and mimic) - ideal for sounds which have no mathematical model (i.e., models that are too complicated or undefined) such as human voices.

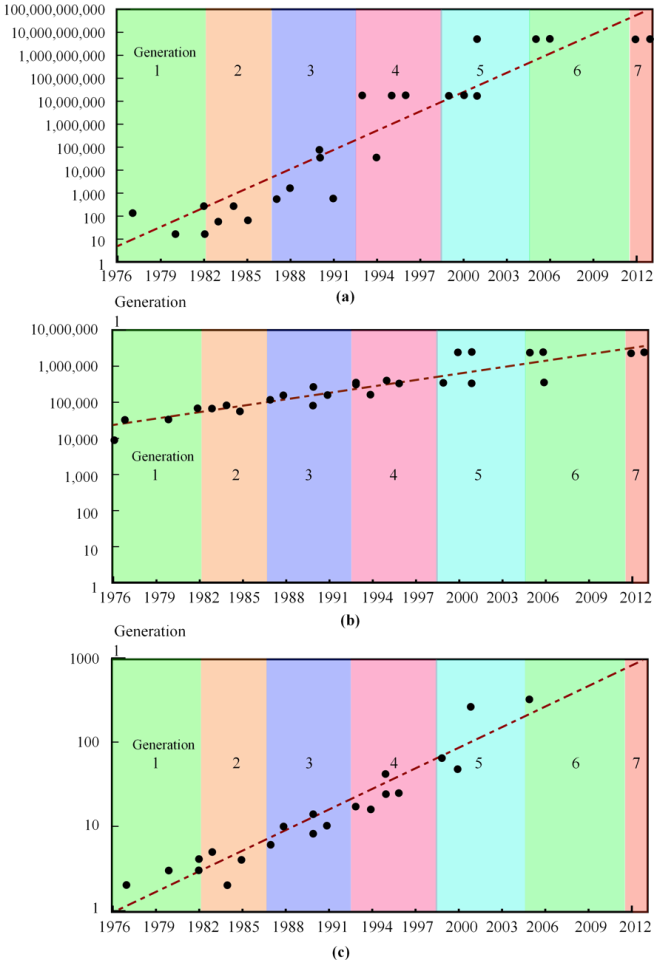

Fig. 2: **Hardware Trends** - (a) Total Colours, (b) Total Resolution and (c) Total Sound Channels. Game console generations and gaming hardware capabilities (from 1976 through 1991, trends in improving quality (sound, resolution and the number of possible colors) - to the point of saturation (i.e., no further improvements are 'noticeable' to the human eye or ear) [19].

5 Sound and Children

Some individuals are more susceptible to sound than others. For instance, video games are a major part of youth and society. A large number of youngsters play video games regularly (see Figure 3), with sound and game play being an extensive part of their entertainment and social development. However, this is a sensitive area of research due to the ethical concerns, risks and dangers of the research to children. There are still huge gaps in knowledge around this topic (especially in video games and in related areas to education, psychological trauma and cognitive development). Of course, sound is a cross-cultural international medium that impacts everyone, not just children, so as sound advancements develop over time they will influence changes in health, education, society and economic legislation. Overall, noises effect children in similar ways to adults (e.g., engagement and cognitive performance) [21], there is evidence that children are more susceptible (lack the ability to block-out/ignore background information such as noises) compared to adults [22].

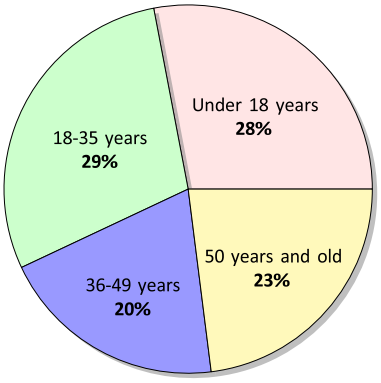

Fig. 3**: Age and Video Games** - Age breakdown of the number of video game players in the United States in 2018 [23].

Sound in Context Immersive technologies are in a state of flux and are rapidly changing, on every level, includes audio contexts - as immersion and realism is a driving force behind today's marketplace (e.g., education, training and entertainment). We also have to remember, sound is not just about realism (playing back real-world noises), it is just as important in the context, type and application of the sound (influences the impact, the mind, the experience and the mood of the individual).

6 Conclusion and Discussion

Sound is an under-appreciated tool in interactive environments (with an assortment of complex underlying mechanisms). Sounds are all around us. From noises in the streets to music on the radio. In immersive interactive environments, sound has the ability to synergistically connect multiple senses leading to a heighten state of engagement and focus. While graphics is the primary technique for engaging and communicating with audiences, in many areas of entertainment, a means to delight and connect, we must not forget the subtle and overlooked importance of `sound' in this experience (e.g., watching a movie without sound).

Sound is just as important in immersive interactive environments as the visual component – and offers limitless opportunities when combined with other senses (exhilarating experiences that capture

the imagination). As we have mentioned in this short article, there are still huge gaps in our understanding of how sound influences our perception, the mind, behaviour and how sound can be used effectively to increase engagement in interactive environments. In summary, we hope after reading this article, the reader appreciates and sees the opportunities and necessity of sound in interactive virtual environments, to say nothing of the need to synergistically combine audio-visual mechanisms to be successful. For instance, sound in video games is paramount, as it tells the player about their environment (giving a sense of presence and time). This connection communicates information in ways visuals alone cannot, while certain combinations of sound and visuals can invoke understanding and emotions that neither could do alone. Finally, we must also note the ability of sound for influencing a user's `perception' – that is, the impact on what a user sees or hears in the absence of information (fill in `gaps' such as when walking down a dark corridors or when lost in the woods) - impacting what the user thinks they hear and see (and vice versa).

REFERENCES

[1] D. Williams and N. Lee, *Emotion in Video Game Soundtracking*. Springer, 2018.

[2] H. Morlock, T. Yando, and K. Nigolean, "Motivation of video game players," *Psychological reports*, vol. 57, no. 1, pp. 247–250, 1985.

[3] Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. Embodied auditory perception: The emotional impact of approaching and receding sound sources. Emotion, 10(2), 216. 2010

[4] R. L. Mott, *Sound Effects: Radio, Television and Film*. Mcfarland, 2014.

[5] W. Cheng, *Sound Play: Video Games and the Musical Imagination*. Oxford University Press, 2014.

[6] K. Collins, "Game sound," *An introduction to the history, theory, and practice of video game music and sound design, Cambridge*, 2008.

[7] T. Tsiatsos, "The impact of background music on an active video game," in *Interactive Mobile Communication Technologies and Learning: Proceedings of the 11th IMCL Conference*, vol. 725. Springer, 2018.

[8] Juslin, Patrik N., and Daniel Västfjäll. "Emotional responses to music: The need to consider underlying mechanisms." Behavioral and brain sciences 31.5: 559-575. 2008.

[9] H. Gemba and K. Sasaki, "Changes in cortical field potentials associated with learning processes of audio-initiated hand movements in monkeys," *Experimental brain research*, vol. 70, no. 1, pp. 43–49, 1988.

[10] C. Spence, "Sound: The forgotten flavor sense," in *Multisensory Flavor Perception*. Elsevier, 2016, pp. 81–105.

[11] B. Bolmont, P. Gangloff, A. Vouriot, and P. P. Perrin, "Mood states and anxiety influence abilities to maintain balance control in healthy human subjects," *Neuroscience letters*, vol. 329, no. 1, pp. 96–100, 2002.

[12] C. Tremblay, F. Champoux, P. Voss, B. A. Bacon, F. Lepore, and H. Theoret, "Speech and non-speech audio-visual illusions: A devel-´ opmental study," *PloS one*, vol. 2, no. 8, p. e742, 2007.

[13] I. Cakir, "The use of video as an audio-visual material in foreign language teaching classroom." *Turkish Online Journal of Educational Technology-TOJET*, vol. 5, no. 4, pp. 67–72, 2006.

[14] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, no. 5588, p. 746, 1976.

[15] D. Deutsch, "The tritone paradox: An influence of language on music perception," *Music Perception: An Interdisciplinary Journal*, vol. 8, no. 4, pp. 335–347, 1991.

[16] J. J. Morgan, "The effect of sound distraction upon memory," *The American Journal of Psychology*, pp. 191–208, 1917.

[17] S. Kalyuga, "When using sound with a text or picture is not beneficial for learning," *Australasian Journal of Educational Technology*, vol. 16, no. 2, 2000.

[18] M. Zembylas and P. Michaelides, "The sound of silence in pedagogy," *Educational Theory*, vol. 54, no. 2, pp. 193–210, 2004.

[19] K. Orland, "Does the power of today's consoles keep up with historical trends?" *Ars Technica*. Available Online. (Accessed 10/09/2019). Nov 2013

[20] A. Farnell, "An introduction to procedural audio and its application in computer games," in *Audio mostly conference*, vol. 23, 2007.

[21] M. Klatte, K. Bergstrom, and T. Lachmann, "Does noise affect¨ learning? A short review on noise effects on cognitive performance in children," *frontiers in Psychology*, vol. 4, p. 578, 2013.

[22] B. Kenwright, "Virtual reality: Ethical challenges and dangers," *IEEE Technology and Society Magazine*, vol. 37, no. 4, pp. 20–25, 2018.

[23] Statista, "Age breakdown of video game players," Available *Online* (Accessed 09/09/2019). 2018.

[24] R. Schroeder, "Possible worlds: The social dynamic of virtual reality technology". Routledge, 1996.

[25] Avdeeff, Melissa. "Artificial Intelligence & Popular Music: SKYGGE, Flow Machines, and the Audio Uncanny Valley." Arts. Vol. 8. No. 4. Multidisciplinary Digital Publishing Institute, 2019.