Acta Ludologica

2024, *Vol. 7, No. 1*

Understanding Auditory Space in Digital Games for Visually Impaired People

Tomáš Farkaš

Mgr. Tomáš Farkaš, PhD.
University of Ss. Cyril and Methodius in Trnava
Faculty of Mass Media Communication
Námestie J. Herdu 2
917 01 Trnava
SLOVAK REPUBLIC
tomas.farkas@ucm.sk

Tomáš Farkaš has been employed as an assistant professor for over fifteen years. His primary research interests lie in the fields of sound design, music creation for digital games, and sound implementation in the context of game engines (mainly Unity 3D, Unreal, and FMOD). He is currently engaged in several game projects as a sound designer, including the creation of two audio games for the visually impaired. His work also encompasses educational games. In addition to his academic pursuits, he is engaged in the recording of audio short stories and podcasts, field recording, and, when time permits, gaming. During his doctoral studies, he also undertook research into the horror genre and its auditory elements.

ABSTRACT:
The article seeks to establish a foundational framework for comprehending crucial definitions and concepts related to auditory space in digital gaming, particularly focusing on audio games designed for visually impaired individuals (VI), often devoid of graphical interfaces. While existing studies often emphasize players' interactions with interfaces and the acquisition of real-life skills, this text explores audio games through the lens of entertainment and immersion. Numerous studies indicate a demand among visually impaired players for more intricate and challenging games, incorporating elements usually used in standard digital gaming experiences. By combining definitions from both audio and traditional digital games, this article broadens the scope of sound design considerations, encompassing various aspects and classifications. It presents several auditory dimensions, consolidating them into a comprehensive dimension called attenuation, putting their roles within a game's context. The author is a sound designer currently working on an audio game; therefore, many of the presented definitions are also a guide that will represent some of his own considerations while working on this game. These include not only the use of binaural or ambisonic sound but also understanding the ways and means of how to work with (3D) space in the context of audio and its possible functions.

KEY WORDS:
ambisonic sound, attenuation, audio games, auditory dimensions, auditory space, binaural sound, digital games, sound, sound design, visually impaired.

DOI:
10.34135/actaludologica.2024-7-1.136-150

Introduction

Audio games have been produced for many years. In fact, the first audio games[1] appeared not that long after the first popular digital games started to emerge in the 70s and 80s. However, many of these used very basic interfaces and were playable mostly only with accessibility software (or, more specifically, text-to-speech software). Since then, a whole genre has started to form, taken different shapes, used different technologies, and focused on different goals. Online platform AudioGames.net describes them simply as games, the main output of which is sound rather than graphics ("Welcome at", n.d.). The primary focus of this article, however, is not the definition and categorization of audio games (although a brief discussion of some of the genres is going to be needed), but rather the definitions of their aspects regarding auditive space and issues that may appear, depending on the type of game and – as we will see later – the players themselves. Whenever audio games are discussed (especially when discussing games created for VI[2]), a very narrow and specific scope is usually utilized, considering the limitations of what can

1 Remark by the author: In this case we mean games that were not primarily created as audio games but could fall under this genre from a current point of view. Many of these include so called *MUDs* – Multi-User Dungeons, since they provide information via text and therefore allow options for access for visually impaired people.

2 Remark by the author: Legally blind or partially impaired.

be done without a graphical interface or more traditional ways of controlling and navigating the game. In this text, we try to expand the definitions, and a more 'general' approach is used, embracing the definitions typical for conventional digital games from the perspective of current sound design theory and practice, as well as because many problems regarding accessibility are irregularly and randomly spread across different players. We believe this approach could bring a new point of view on audio games and extend how they are created. Another goal of this article is to highlight that there is no definitive and unequivocal way to create audio games that can satisfy all visually impaired players. This is due to the same reasons why players without visual disabilities do not play all games but usually enjoy specific genres and choose not to play others.

Among many of the studies, few try to use a similar approach. For example, Andrade et al. (2019) aim to understand the 'lived experience' and practices of players with visual impairment in their interactions with games and play. Their approach presents a more practical look at how VI individuals perceive digital games and their preferences and criticisms. One of the details from their article points to the fact that there is no consistent set of rules or audio cues (a notion we define later) that would be used in different games across all the different audio games, and that players usually must learn a new set of 'auditive language' from the beginning with each new game they play (Andrade et al., 2019). A similar idea was presented by White et al. (2008), stating that the development of a standard for using audio to communicate 3D data in an operationally deterministic manner remains a goal of future work. Although this article was written in 2008, this issue is still relevant. This notion could be contextualized by Bălan et al. (2015b). In their research, they found that information from the central and the peripheral regions of space is processed differently by the auditory system of the VI – in one of the experiments, VI participants and sighted (blindfolded) participants were required to locate the target sound and to specify if there was any difference in its spectral content (the presented sound was 555 Hz pure tone frequency) when more sounds were played consecutively. This study showed that VI were more efficient in detecting and localizing sound targets[3] but less efficient in discriminating the spectral content. Simply put, VI were better at defining *where* the sound is coming from but worse at *what* the sound is or how its spectral character changes.

It should be pointed out, however, that many of the experiments like this are performed while using 'sterile' and somewhat artificial sounds. We do not usually hear pure frequencies in real-life situations. It is widely known among sound designers that many individuals, regardless of their visual ability, may struggle to distinguish between different pure tone frequencies or accurately describe changes in pitch. Some researchers (and game developers) work with white and pink noise. Coincidentally, another piece of research from Bălan et al. (2015a) used a combination of white and pink noise in varying proportions to create a sense of space, more precisely – front and back sound localization. While the choice of sounds like this is understandable from a research perspective, it contrasts with what some of the 'different approach' studies present.

Many of the interviewed VI players expressed a desire for more immersive experiences with the use of 3D audio in games, and even some of the features we know from conventional games like RPGs – for example, different types of armour in the game that would have distinctive sounds according to the material (Andrade et al., 2019).

Urbanek and Güldenpfennig (2019) take a similar approach in their research as outlined in the paragraphs above. They talked with several 'game veterans' of sound-based

3 Remark by the author: Which was a logical assumption, since it's been proven by many studies in the past that visually impaired people usually develop other senses (like hearing, smell, tactile skills etc.) more.

gaming and conducted 14 interviews, creating a set of their own experiences, revolving mainly around creativity, play, and social exchange. The results of these interviews closely resemble what was also found by Andrade et. al., involving the desire to play digital games in general, but also limitations of known audio games and main-stream games as well (in this context, video games, in general, are referred to as main-stream games). Michael Urbanek's body of work focuses on audio games from several different perspectives, for example rethinking prototyping for audio games (Urbanek & Güldenpfennig, 2017), understanding audio games online (Urbanek et al., 2019), or even working towards an online audio game editor (Urbanek et al., 2018b) and creating a set of 'anti rules' for audio game design (Urbanek et al., 2018a).

Some researchers directed their focus at audio games more than 20 years ago. Sánchez and Lumbreras (1999) studied the potential of audio games in the context of supporting learning and mental map generation in visually impaired children. They produced and utilised the game *AudioDoom* (a simplified audio-only version of a popular shooter) to test how well children would be able to reconstruct the map from this game using Lego bricks (Lumbreras & Sánchez, 1999; Sánchez & Lumbreras, 1999).

One of the points of this article is to talk about audio games as 'games' – that means considering more traditional notions like immersion. This notion would not hold against simplistic sounds, strictly using clean sound waves or types of noise. In the following chapters, we will present a combination of definitions that could be applied to audio games for VI, the same as for traditional players, as well as point out some potential problems that might be related. Even though there are currently hundreds of audio games for VI, many of these are still very simple in nature, and although there have been a lot of 'we-live-in-the-golden-era-of-sound' statements in recent years, the industry is still leaning heavily towards visual perfection, VR/AR experiences, leaving sound somewhat behind.

Another significant area that our article only slightly touches upon is that of accessibility. Accessibility options are slowly being introduced in the mainstream consoles. Microsoft and Sony, with their recent models of consoles, began to implement ways for people with different disabilities to be able to interact with the consoles themselves, and in a limited fashion with the games. Visually impaired users (both legally blind and partially impaired) now have the option to use text readers in the menus and some of the games, as well as zooming the screen, using high contrast modes and colourblind mode, along with other functions, which was not so common just a few years ago. For example, Xbox launched the Microsoft Game Accessibility Testing Service in February 2021, a program developed for publishers and developers to validate the accessibility of their games (Mortaloni, 2023), following their release of the Xbox Adaptive Controller, released in September 2018 (Wilson, 2019). Sony followed this with their version of an adaptive and customizable controller called Access, with the release of Project Leonardo in January 2023 (Nishino, 2023). It might seem unrelated to the audio games topic, however, many VI players expressed issues with actual controls – for most players, WASD keys as keys for movement are standard, but VI players expressed the need for regular arrow keys, as well as problems with navigating games with a keyboard and mouse, which could be an interesting idea for separate research.

The previous paragraph suggests an increasing trend in taking more prominent steps toward players with disabilities. However, this also means that more time must be spent studying how to implement different kinds of sounds in different games. In the next chapter, we discuss some of the most important definitions regarding audio games and put them in the context of traditional sound design.

How to Talk about Auditory Space

We can approach audio in digital games from several points of view. Some are more technical, and some bear an aesthetic element, making them more ambiguous. All of them, however, can be applied to audio games in one of their forms and functions. Functions will also be discussed since they are essential to any digital game development. Some of the definitions will also be put in the context of sound design decisions, because the fact that something *can* be used does not necessarily mean it *should* be used – and vice versa.

Before we begin to discuss the various auditory dimensions in more detail, it is necessary to define the concept of auditory space, which we have already mentioned several times in our text. This notion does not exist as a dictionary term expressing specific features. Rather, it represents a set of different elements, properties and ways in which space in digital games can be thought of in the context of their genre and type, but also more broadly in a technological and therefore historical context. Auditory elements in such conjunction then represent the way in which sound designers respond to the specific situations and needs of a particular game. However, it is important to be aware of the functions of sound and their possible uses. Murphy and Neff (2011) say that no matter what approach is employed in creating sound design, the primary objective is always the enhancement of the virtual scene and the enrichment of the user experience, which correlates with the opinions of the majority of other authors involved in sound design theory and practice.

Based on this rough definition, we can then approach auditory space from several possible angles. Very simplistically, we can think about what (and if) the player hears, in what part of the game (or from where), how they hear it, and most importantly, why they hear it. When we construct the definition of auditory space in this way, it becomes clear that the genre and type of a particular digital game are very relevant. The way auditory elements are perceived in a 2D game (whether designed in a top-down style or a platformer we watch from a side-view) is different in a 3D game. In the latter case, for example, it is essential to consider the positioning of the so-called 'listener', as the response of sound to player actions differs between games that use first-person view and third-person view.

The historical and technological context is then defined by the possibilities of auditory space itself, as the potential possibilities of working with and implementing sound evolve with each passing year.

The Definitions of Auditive Elements in the Context of Audio Games

The easiest way to begin the definitional part of this text is probably defining so-called *auditory dimensions*. This term is usually used to describe three general characteristics of sound – loudness, pitch, and timbre. We will expand these since audio games are an interactive medium requiring an extended approach. Bălan et al. (2014) add directionality, distance, and externalization to the previous three. In another research (Wang et al., 2022), auditory dimensions are defined as pitch, volume, panning, length, tapping, and timbre.

Although this research did not focus on audio games, but on the sonification of different data types, the ‘tapping’ dimension is a very interesting concept for our topic. For our purposes, we will combine the presented auditory dimensions and talk about *volume* (or *loudness*),[4] *pitch*, *timbre*, *directionality* (which can be understood as a synonym for panning and positioning), *length* and *repetition* (again a synonym for tapping, but more suitable for us). We add one more category, however – *attenuation*, which is technically a combination of more elements. However, in the context of sound design theory and practice, it is one of the elements of the ‘mix’, a fourth category of sound as defined by Bridgett (2021).

Volume (or loudness) can be understood as one of the most fundamental concepts in audio in general. In the context of audio games, however, it can be one of the most effective ways to communicate distance (whether from any object, objective, or location). It also can carry the meaning of ‘amount’ in general. Suppose there is a game in which the important information for the player cannot be conveyed by a voice or any other means. In that case, the volume of specific sounds can represent a level of anything that needs to be understood – remaining life with the increasing sound of a heartbeat, the height of our position on the map with the increasing sound of the wind, and many other things. Volume can also be understood from the perspective of silence. If something dramatically lowers in volume or completely stops playing, it is usually a sign to the player that she or he should start paying attention. Volume in connection with music can also bring some surprising effects – Kellaris et al. (1996) suggest that very loud music can make time feel perceived more slowly by its listeners.

The *pitch* of the sound is more complex than the volume. It can be simply explained as a way to perceive sound frequency. It is usually measured in Hz, and most people can distinguish between 20 Hz to 18 Hz.[5] Some studies (e.g. Wang et al., 2022) proved that pitch is one of the most effective ways to communicate information. From the sound perspective, it can be easily explained: when the pitch of the sound is changing, our ears usually perceive it quite significantly. The pitch can be translated to ‘tones’ from the musical perspective. When we imagine a piano, lower tones represent a lower pitch, and the higher we go, the higher the pitch. This example can also demonstrate how complicated using different pitches can be. The tones (or the frequency) of the sounds can be easily distinguishable when placed a little further apart. However, if we were to play two keys that are only one semitone apart, many people without any musical training would have problems identifying ‘what’ is happening with the sound and would not be able to describe whether tones are played ‘higher’ or ‘lower’. The same applies to two similar frequencies of pure tones. Pitch as a concept or a game mechanic can be assigned to different things. Some audio games use changes in pitch precisely to tell the player whether a sound is located at the front or the back – which is, by the way, one of the most common problems in this genre since a 100% realistic representation of sound localization is not possible, even with the use of the most advanced binaural or ambisonic microphones, partly because individual HRTFs (Head Related Transfer Functions) demand complicated measuring techniques (see Farkaš, 2018). To complicate things further, the pitch is a feature all sounds have. That means if we change the pitch of the voice, we will get completely different effects than if we change the pitch to, say, the sound of an elevator. We can create a chipmunk or a Darth Vader voice, but when we gradually raise the pitch of an elevator sound, suddenly, we have created a charging intergalactic laser beam.

4 Remark by the author: However, it should be mentioned that loudness in this regard does not represent subjective temporal perception of sound pressure in contrast with volume, which is measured a little differently – as defined by acoustics, mixing or mastering process. We use volume and loudness as synonyms here to simplify the definitions.

5 Remark by the author: This is a very approximate and general definition. Younger people are able to hear more frequencies and as we get older, our ability to hear higher frequencies (above 16 Hz and more) usually decreases. On the other end, we tend to ‘feel’ lower frequencies (under 20 Hz) more than ‘hear’ them.

Timbre can be understood as the quality of the sound, defined by its overtones, varying in their frequencies related to fundamental tone, volume, and waveform patterns (Savage, 2014). Simply put, timbre represents a 'colour' perception of the sound. In connection with pitch (and volume), for example, we can see how complex the sound can get – two different people singing in the same pitch and volume can sound drastically different only because their vocal cords grew differently. The same applies to musical instruments – timbre distinguishes all musical instruments from each other. We can usually tell if we hear a Spanish acoustic guitar, a dulcimer, or a harp (all three use strings), and most people would probably be able to distinguish between a trumpet, a flute, and a trombone.[6] The problems start to occur when we want to use atonal sounds[7] – or *Foley*. Foley represents many things and depends on the viewpoint from which we define it. From the traditional movie perspective, Foley is the art form that adds believable sound effects to on-screen character movements (Marks, 2017). In sound design theory and practice, Foley may represent many more things than just character movement. Some sound designers instead use the term 'sound effects' (although this could cause a little confusion as well),[8] representing all the sounds made by objects around – the sound of a door, table, water, simply all the things that make sounds in one way or another. The element of timbre can be problematic when we put it in relation to the 'where vs. what' idea. The sound of some objects or sound effects would be recognized easily by VI and sighted people (e.g. car engine, plane, singing of birds, typing on a keyboard, or breaking glass), but many other sounds could cause problems. For example, a creaking wood sound could represent a rocking chair, a wooden window shutter, or an old dry tree. Without any visual representation or any other 'help', some sounds are simply too difficult to refer to a single thing, even for sighted people. That means timbre as an auditory dimension can be a mighty but double-edged sword and should be considered carefully, mainly in connection with sound functions, discussed later.

Directionality, or panning/positioning as an auditory dimension, represents sound elements that rely heavily on implementation decisions. The implementation process constitutes the methods of how the different sounds are put into a game. In the context of directionality, several points of view should be considered if we want to stay within this article. The most basic approach is the concept of stereo panning. In a traditional stereo field, we can place and move the sound to the left or right. If combined with a volume dimension, this is probably the easiest way to create a perception of space in an audio game. This, however, is effective only if our gaming space is 180 degrees in front of the player. Another step would be combining the previous two dimensions with the pitch – all the sounds 'behind' the player could be gradually turned down a pitch to signal the change of position, making the player able to perceive a 360-degree area around them. Again – this works in theory but should be carefully tested before implementing in a real game. Nevertheless, directionality can also mean more specific things. When, for example, implementing sound created to be played constantly (loop), we can 'place' it anywhere in the location of our gaming space. Game engines like Unreal or Unity usually have several options that can be tweaked to change how the sound is perceived. Several parameters should capture our attention if we want to implement the sound in 3D space. Every sound can have an inner and outer circle of volume falloff, and – what is more important – other

6 Remark by the author: Of course, just as with the example with pitch, some people would have difficulties distinguishing viola, violin or maybe even high tones played by cello.

7 Remark by the author: In this regard, we do not mean atonal music as compositions written in the early 20th century, but more broadly sounds without clearly distinguishable tones or pitches.

8 Remark by the author: In the context of DAW (Digital Audio Workstations), sound effects could be mistaken for audio-effects, which refer to different ways of how to change/manipulate sound quality (like equalization, compression, flanging, chorus, pitch modulation etc.).

spatialization parameters change how this sound reacts if we go closer to it and pan our camera (or listener) left/right. The audio listener allows us to hear the audio in the game and represents the auditory perspective rendered when playing spatial audio. Usually, this component is added to a game camera (Sinclair, 2020) – in the environment of an audio game, the audio listener is probably one of the most complex terms, simply because there is usually no camera involved, and our listening experience is solely based on the game type. This can mean anything from a fully 3D environment to a 2D top-down view. Spatialization settings can quite significantly change our hearing experience and immersion. If we set this parameter too low, sounds can quickly and abruptly 'jump' from the left to the right channel, which might break the immersion and confuse the player in the case of an audio game. If we set the parameter too high, left vs. right information can only be perceived from a greater distance since when we are closer to the source, we will simply hear it in both channels. In middleware like FMOD or Wwise, built-in parameters can help us with the directionality of a sound. The elevation parameter is useful if we want to change our audio source on the vertical distance between the source and a listener, and can create a simple vertical occlusion effect, respectively, the impression that something is up or down. There is also an event cone parameter in FMOD, tracking the position and direction of an audio source relative to the audio listener – this way, we can, for example, simulate a situation when we face a talking NPC whose voice gets quieter or muffled when we go around or behind it.

Length of a sound is a self-explanatory term, but it can also contain more information than it might seem. Short sounds can be used differently than longer sounds, and there is also an 'infinite' length if we use looping sounds. In combination with simple sound effects,[9] short duration can catch our attention if used carefully, and with correctly set directional settings, volume, and pitch, this kind of sound can (and usually is) be used as a navigational beacon for a player with visual impairment.[10] Some of these are used as so-called attractors – they can actively draw player's attention for different reasons. Loops are usually used as a wider texture, describing the location or mood of a certain game area. The shorter the sound, the more obvious it can be, but even a 5-6-second-long sound can be considered short in this manner. A few other terms should be defined in connection with the length of sounds and beforementioned attractors. First, so-called *auditory icons* represent short, icon-like sound events that have semantic connections to the physical events they represent (Csapó & Wersényi, 2013). They are easy to interpret and learn, and in the context of an audio game, they can represent short sounds associated with specific objects, helping the players to identify what is happening around them, what kind of material they are walking on, and similar details. Second, there are so-called 'earcons', which can be understood as message-like sounds (for example, consecutive short notes) that gain meaning through abstract relationships between signifier and signified. Users (or players in our context) are required to explicitly learn how earcons are linked with different events (Csapó & Wersényi, 2013). The third term to consider is the so-called 'spearcon'. This is an interesting concept because, according to past research (see Jeon & Walker, 2009), spearcons were found to be superior to auditory icons or earcons when tested in different applications. They represent sounds obtained by speeding up speech, usually to the point that they are no longer recognizable as speech but preserving their original

9 Remark by the author: And here we can talk literally about synthesized beeps and bloops, as well as about more realistic sounds.

10 Remark by the author: Although these kinds of sounds are also used in traditional games for sighted players.

pitch[11] (Csapó & Wersényi, 2013). According to Juan and Walker, they resemble fingerprints since the acoustic relation between these and the original speech phrases is still present.

Repetition is a concept that is not a literal auditive dimension, but which might paradoxically be one of the most important. Similar to volume and pitch, the repetition of sound can easily convey many types of information, especially in combination with short sounds. If we correctly use it and teach the player at the very beginning (and of course, again, if we combine it with, say, the changes in the pitch), we can use different numbers of repetitions to bear information concerning things like correct/wrong, open/close, go/stop, attack/run, the worth of a treasure, durability of a weapon/armour, or virtually any other thing that can be assigned with short consecutive sounds. In an extreme case, we could navigate the map without any graphical interface using just one short sound representing the north/south/ /east/west direction and one sound that would play if we walked into the wall. One 'beep' could be north, two beeps west, etc. Another thing to consider is something we briefly mentioned previously – VI usually develop other senses more, which means they can distinguish shorter consecutive sounds better, even if played very quickly. This can be demonstrated by the fact that text-to-speech generators implemented into some audio games usually read the text extremely quickly, sometimes making it problematic for sighted people to understand. Regarding the length, repetition should be considered when assigning sounds to important objects like enemies. If there is a longer pause between enemies' 'barks'[12], it can be problematic for a player to localize them successfully.

Attenuation is our own way of combining most of the auditive dimensions into one complex entity, and can be understood as a governing element over all the above. We have already slightly touched on some of the methods attenuation presents when describing pitch or directionality. However, this term can refer to more ways to process sound and can become very complex. Bridgett (2021) defines attenuation as setting up the amount of volume, low pass, or high pass filtering that occurs on a 3D object's sound as the listener in the game gets further away from the sound source. Usually, attenuation is created using so-called *attenuation curves* – these can be assigned to any parameter or audio effect we can imagine. Curves can have different shapes, and given parameters react accordingly. The linear curve would probably be used in different scenarios than logarithmic, and the game engines also offer custom shapes. Attenuation can include most of the auditory dimensions in one complex combination. For example, suppose we want to create a sense of space for an object in a room. In that case, we can set up curves like this: one curve is going to affect volume according to the distance from the sound source, a second curve is going to represent the amount of reverb added to the sound (this has to be applied in reverse because we want to hear more reverb the further away from the object we are), the third is going to cut off high frequencies if we leave a certain distance or a room itself and if we want to add details, we can add another layer of sound, audible only if we move very close to the object. All of this will be created with the set of multiple attenuation curves, but – this method allows us to process all of these elements with just one adaptive parameter linked to distance. If we have a capable programmer in a team, we can create a simple attenuation that makes all the sounds 'behind' us sound slightly more pitched

11 Remark by the author: This technique is easily accessible in most standard DAWs as so-called time-stretching, which can have different results when using different algorithms. If we speed-up (or slow down) the time in which the sound is played, we have an option to either preserve its pitch, or to produce more natural sound, that would create a chipmunk-like effect or completely slowed-down low-frequency heavy sound, similar to if we slowed down a playing vinyl record.

12 Remark by the author: Some sound designers use this term for short sounds or one-liners of NPCs or enemies, functioning as their sound 'identity', strengthening immersion and sometimes even offering information.

down, filtered, or with reduced mid-content.[13] Attenuation should be considered carefully when dealing with important parts of a sound design in any game. If a story is an important element of our concept and is conveyed by a voice, we want to hear it from a greater distance. That means that the attenuation curve of the volume of this voice is probably going to be longer than that of other, less critical sounds, and its shape is not going to be linear but rather convex.

As we can see, all the auditory dimensions (including attenuation) create a sophisticated and complex system of variables. This perspective makes sound design a wildly extensive area in which the most straightforward solution using just a few sounds can resolve in functional gameplay and intelligible game navigation. Likewise, we can create complex and immersive soundscapes that communicate information effectively. There are other dimensions that we could analyze in this manner next to volume, pitch, or directionality. One such could be *movement*. So far, we have been talking about sounds that do not move, but that is not always the case in games (it is, rather, the opposite). Again, this dimension could be viewed from the perspective of understanding vital information in any digital game. If a short, simple sound travels from the left to right channel, it may have a different meaning than if it travels in the opposite direction. If a sound is added to a moving game object as a component, a player can perceive the changes in its location, distance, etc. This means that movement as a dimension of sound could be used in many categories, whether it represents the sounds of a user interface, game objects, or enemy location.

Auditory Dimensions in the Context of Functions

The functions of sounds in digital games are a vast and expansive topic, complicated by the fact that every game is different. Even more so if we want to discuss audio games without visuals. Functions as such could be defined in many ways, and some authors offer a simple list of these, from which we can 'bounce' further in defining more specific terms. Zdanowicz and Bambrick (2019) offer six primary functions of sound: providing sonic feedback to the player, communicating emotion/setting the mood, providing a sense of space, defining realism through literal and non-literal sounds, establishing a sonic identity or brand, and establishing structure, narrative, and pacing. However, they further claim that this is not an exhaustive list but rather a framework of important notions that overlap in many ways. Bridgett (2021) does not define the functions at all – but instead offers a large number of examples from which we can easily read all the various functions and methods of working with sound. Sinclair (2020) creates three broad function categories – inform, entertain, and immerse, discussing them from the perspective of the purpose of audio in games in general, further dividing these into smaller, more detailed categories (like, e.g. how to inform and what to inform about). From our perspective, it is easier to ask several simple questions. For example, what game genre are we dealing with, and how and why can a sound communicate it? What types of sound do we need in this game? Where, when, and why are they played? Do we need these sounds to navigate the game, or are they just creating a more immersive experience?

13 Remark by the author: Mid content means a part of an audio that is not different between left and right channels and is basically a mono content in the 'middle'.

The most crucial thing in this context is understanding audio from the perspective of all the different dimensions we discussed before, and asking one specific question all the time: What does the game really need? This question is probably the most important idea that many sound designers and authors proposed and can be extended to *what a player really needs to hear* to successfully play and navigate the game. Of course, the audio game genre complicates this a bit further because many sound functions in digital games are linked to visuals. Considering all the obstacles VI players can face, we can get around this by incorporating other methods that would function similarly. A synergy effect usually occurs when we combine more than one way the sound exists within the game. One area we have not discussed so far in this text is *haptic feedback*. Regarding console controllers, and even more so with smartphone devices, haptic feedback can be efficiently executed by the vibration capabilities of these devices. Regarding some of the basic functions of sound, vibration can be performed as an extension. An information category can be an interesting example of how sound can be enhanced with vibration to provide further details and guidance for a player. Suppose a game needs to communicate to a player that there is a monster behind her or him, in addition to attenuated sounds of footsteps (which could still confuse localization if not appropriately filtered). In that case, a vibration can be added to these steps – but only when the monster is behind (e.g. anywhere else than in the 180-degree field in front of the player). Steps treated like this also provide feedback and can even increase the level of emotions the player feels at that moment since the level of excitement or fear rises. It should be pointed out, however, that the topic of haptic feedback covers a large area of its own and its examination is not the primary goal of this article.

The whole point of sound functions in any (audio) game is simple – they should increase immersion and provide a player with enough information to successfully play, navigate, understand, and finish the game. The way we understand this is solely based on the specific game and after asking two questions stated in a previous paragraph. If we 'teach' a player at the very beginning of the game that when harp sounds are played, she or he is in a safe zone/merchant area, she or he will remember it until the end of the game. Likewise, associating the sound of drums with enemies will make her or him more aware of the surroundings later in the game when the sound of drumming starts (we can even divide enemy strength by types of drum if we want). Many more examples like this could be mentioned. If we wanted to inform the player about her or his armour status, several versions of armour hit (or a movement, if not in a battle) could indicate serious injuries or near death (one with the prominent sounds of metal clinging could represent full health and safety, one with more 'broken metal' elements and more low frequencies could stand for medium injuries, and the sounds with elements of a human voice, saying different versions of 'ouch!' or 'aah!' would signify danger of dying). A similar approach could be used in a puzzle or an adventure game. Sound (and music as well) can tell a player if she or he is progressing towards a successful finish of a puzzle or, on the contrary, whether they 'messed up' and need to take a few steps back.

Another vital question to be asked is whether (or, more precisely, 'why' and 'how') we should talk about space in audio games. For sighted people, the perception of so-called 3D space is natural. For VI, three-dimensional space means something completely different. If we want to create an immersive '3D' audio experience, a 2D map navigated from a top-down view can be generated, with all the sounds implemented as positional 3D audio with different directionality, spatialization, and falloff settings – and this kind of game design is still perceived with 3D depth. Elevation elements could still be 'baked'[14]

14 Remark by the author: This term refers to rendering or exporting of the sound. Even if we export audio file as a linear audio or loop, we can use different plugins in our DAW of choice to render some of the sound elements or layers positioned 'up' or 'down'. There is a plenty of free and paid plugins that are able to create a perception of elevation like Sennheiser DearVR MICRO (or DearVR PRO respectively), Soundfield by Rode, or standalone software like SPAT Revolution.

into positional sounds, and even without the possibility of looking up or down, the illusion of space would be constructed. In addition, the 'proper' elevation can still be created only with advanced sound systems like Dolby Atmos with some of the speakers placed above the audience/player, which, of course, is hard, if not impossible to achieve with most audio games played on headphones. Different methods of creating an illusion of sounds being localized in 3D space (especially at the back, up, and down) will be the subject of our following research. We will test the sound perception of this kind produced with the help of binaural/ambisonic VST plugins using standard stereo/mono audio files, as well as with an ambisonic microphone.

Examples of the Application of Different Attenuation Methods in Real Projects and Potential Drawbacks

As previously mentioned, the author of this article is also a lead sound designer and works on several projects related to audio games. One of these games (*Via Echo*, scheduled for release on Google Play in the first half of 2024, developed by Blind Octopus Studio) utilizes several of the attenuation methods mentioned earlier. The main mechanics of the game are based on camera rotation while the player remains static. The primary goal is searching specific (sound) objects in 360-degree field around the player, which use a very narrow sound directionality making them difficult to find. After correctly locating the sound, the player taps the screen, triggering another mechanic associated with 'catching' or 'shooting' sound. The primary issue that required resolution was the communication of the front versus back sounds location. We used several attenuation curves filtering different aspects of sound. The main difference between default processing and our game was switching FMOD's default audio listener for a third-party plugin called Resonance Audio. This listener allows for easier and more precise adjustment of sound's stereo spread, as well as its directionality. One of the benefits of this plugin is its built-in ability to filter sounds more realistically even on default settings, allowing the player to locate sounds (or distinguish their direction) more precisely (see Farkaš & Schwarz, 2023).

To achieve a complete auditory immersion, we have created four distinct ambient atmospheres for each level. These atmospheres are looped and randomized and are placed around the player. We have ensured that the spread of these atmospheres is wide enough to create an immersive experience but not too wide that the player loses the ability to perceive their movement in headphones or stereo field. We have also applied specific filtering and attenuation to aid the player in locating sounds, thus avoiding any difficulties. Each ambient atmosphere has therefore attenuation curves linked to 'direction' parameter in FMOD. When the camera is rotated, sounds are not only panned from left to right, but they also change volume attenuation (up to -5 dB when behind the player), low-pass filtering (reducing frequencies to 6kHz when behind the player), simple reverb (gradually increasing when the camera moves away from a specific atmosphere), and a gain of low frequencies in the reverb (gradually increasing when the camera moves away), adding another layer of depth to the sound. As each level contains four sounds of this type, a simple camera rotation affects 4x4 attenuation curves at any given moment. This is combined with the 'whole' sound rotating 'around the head' of the player. This kind of attenuation makes it easier to navigate the level and communicate front versus back information to players. The result is an immersive auditory experience for the player while still retaining the ability to communicate the direction of the camera. Another detail to consider is the use of 'snapshots' in FMOD. These are set up so that every time the player interacts with the 'catching' or 'shooting' mechanics, all ambient sounds are lowered by several decibels (a technique known as 'ducking' in sound design). This highlights the interactive sound and makes it more audible. It is worth noting that the camera is locked in a horizontal position. *Via Echo* is currently in the testing and polishing phase, and feedback is being provided by both sighted and visually impaired players.

The same studio is currently developing *Via Memories*, a digital audio game that serves as the logical sequel (not in the sense of narrative aspects) to *Via Echo*. While it will utilize some of the mechanics from its predecessor, there is one major difference: players will be able to move freely in certain areas. This will introduce new challenges, as controlling a static camera in relation to attenuation is less problematic than free movement. A potential solution to this issue is to utilize a technique employed in a previous game. In *Via Echo*, all sound objects located around the player are programmed to spawn facing the player. This is crucial in the context of Resonance Audio's listener, which is always linked to a specific axis in the Unity3D engine. Implementing this method could aid in navigating the game with free movement, as most of the attenuation curves should function similarly, regardless of whether the object is stationary or in motion.[15] Setting the correct attenuation settings should be relatively accurate even if the object is moving while still facing the player. Additionally, communication of the borders or edges of free-movement areas will be an important topic in the upcoming game. However, this requires further discussion and research.

Conclusion

What does the audio game need? This article presents a simple framework of ideas to be understood before developing a space-oriented game for visually impaired (VI) players. It is important to recognize the auditory dimensions and sound functions that can immensely change how a game is played and perceived. As we have shown in our article, it is crucial to understand auditory space in digital games in general. Genre and type of game heavily affect this field and create a complex body of different elements that need to be understood in all possible combinations, resulting in the need to ask very specific questions in the process of creating any given game. The genre of audio games complicates this notion even more. The lack of standard visual elements and the sole reliance on auditory elements makes this process complicated and unpredictable, even after several decades of audio game history, as some of the recent studies (specifically those working with VI players) suggest.

Our own projects utilize many of the methods discussed in this article. In the coming months, we will be developing a story-based audio game with immersive sound design. Our initial prototype has already undergone testing with visually impaired players. The results of these tests will be the subject of future research. However, we plan to conduct further testing using various methods in the coming months. It is crucial to collaborate

15 Remark by the author: With that said, we are still talking about a horizontally locked camera. The ability to move camera up or down brings a new set of issues.

with visually impaired players and listen to their needs. The perception of auditory space for VI individuals differs significantly from that of sighted individuals. Using a blindfold alone is insufficient.

Additionally, it is important to consider whether control methods designed by sighted individuals, such as tapping, swiping, or specific finger movements, will be equally effective for visually impaired players. An immersive experience can be created without a complicated 3D environment. By considering key game design elements, their combination, and methods of implementing sounds and navigation, surprising results can be achieved. This is especially true when we consider the possibilities of modern game engines and the increasing power of smartphones, computers, and other devices. Audio games can introduce visually impaired individuals to the expanding world of virtual worlds and social experiences, an opportunity that is still alarmingly dormant.

BIBLIOGRAPHY

Andrade, R., Rogerson, M. J., Waycott, J., Baker, S., & Vetere, F. (2019). Playing blind: Revealing the world of gamers with visual impairment. In S. Brewster, & G. Fitzpatrick (Eds.), *CHI'19: Proceedings of the 2019 CHI Conference on human factors in computing systems* (pp. 1-14). ACM. https://doi.org/10.1145/3290605.3300346

Bălan, O., Moldoveanu, A., & Moldoveanu, F. (2015a). Binaural sound analysis and spatial localisation for the visually impaired people. In K. Blashki, & Y. Xiao (Eds.), *9th Multi conference on computer science and information systems 2015 – interfaces and human computer interaction* (pp. 331-335). International Association for Development of the Information Society.

Bălan, O., Moldoveanu, A., & Moldoveanu, F. (2015b). Spatial auditory representation in the case of visually impaired people. *Romanian economic business Review, 9*(1), 1-11.

Bălan, O., Moldoveanu, A., Moldoveanu, F., & Dascalu, M. (2014). Navigational 3D audio-based game-training towards rich auditory spatial representation of the environment. In R. Lozano (Ed.), *2014 18th International conference on system theory, control and computing (ICSTCC)* (pp. 682-687). IEEE. https://doi.org/10.1109/ICSTCC.2014.6982496

Bridgett, R. (2021). *Leading with sound: Proactive sound practices in video game development*. Focal Press. https://doi.org/10.4324/9781003082521

Csapó, A., & Wersényi, G. (2013). Overview of auditory representations in human-machine interfaces. *ACM Computing Surveys (CSUR)*, *46*(2), 1-23. https://doi.org/10.1145/2543581.2543586

Farkaš, T. (2018). Binaural and ambisonic sound as the future standard of digital games. *Acta Ludologica*, *1*(2), 34-47.

Farkaš, T., & Schwarz, M. (2023). Using A. I. for the generation of visual references in digital audio games from the perspective of sound design. In M. Hossová Prostináková, M. Martovič, & M. Solík (Eds.), *Marketing Identity: AI the future of today* (pp. 61-79). UCM. https://doi.org/10.34135/mmidentity-2023-06

Jeon, M., & Walker, B. N. (2009). "Spindex": Accelerated initial speech sounds improve navigation performance in auditory menus. *Proceedings of the human factors and ergonomics society annual meeting*, *53*(17), 1081-1085. https://doi.org/10.1177/154193120905301708

Kellaris, J. J., Mantel, S. P., & Altsech, M. B. (1996). Decibels, disposition, and duration: The impact of musical loudness and internal states on time perceptions. In K. P. Corfman, & J. G. Lynch (Eds.), *Advances in consumer research* (pp. 498-503). Association for Consumer Research.

Lumbreras, M., & Sánchez, J. (1999). 3D aural interactive hyperstories for blind children. *International Journal of Virtual Reality*, *4*(1), 18-26. https://doi.org/10.20870/IJVR.1999.4.1.2661

Marks, A. (2017). *Complete guide to game audio. For composers, sound designers, musicians and game developers*. CRC Press. https://doi.org/10.1201/9781315758442

Mortaloni, A. (2023, May 18). *Xbox celebrates commitment to accessibility on global accessibility awareness day 2023*. Xbox News. https://news.xbox.com/en-us/2023/05/18/xbox-celebrates-commitment-to-accessibility/

Murphy, D. & Neff, F. (2011). Spatial sound for computer games and virtual reality. In M. Grimshaw (Ed.), *Game sound technology and player interaction: Concepts and developments* (pp. 287-312). IGI Global. https://doi.org/10.4018/978-1-61692-828-5.ch014

Nishino, H. (2023, January 4). *Introducing project Leonardo for PlayStation 5, a highly customizable accessibility controller kit*. PlayStation Blog. https://blog.playstation.com/2023/01/04/introducing-project-leonardo-for-playstation-5-a-highly-customizable-accessibility-controller-kit/

Savage, S. (2014). *Mixing and mastering in the box. The guide to making great mixes & final masters on your computer*. Oxford University Press.

Sánchez, J., & Lumbreras, M. (1999). Virtual environment interaction through 3D audio by blind children. *CyberPsychology & Behavior*, *2*(2), 101-111. https://doi.org/10.1089/cpb.1999.2.101

Sinclair, J. L. (2020). *Principles of game audio and sound design. Sound design and audio implementation for interactive and immersive media*. Routledge. https://doi.org/10.4324/9781315184432

Urbanek, M., Fikar, P., & Güldenpfennig, F. (2018a). About the sound of bananas – Anti rules for audio game design. In J. L. Vilaça, T. Grechenig, D. Duque, N. Rodrigues, & N. Dias (Eds.), *2018 IEEE 6th international conference on serious games and applications for health* (pp. 1-7). IEEE. https://doi.org/10.1109/SeGAH.2018.8401361

Urbanek, M., & Güldenpfennig, F. (2017). Rethinking prototyping for audio games: On different modalities in the prototyping process. In L. Hall, T. Flint, S. O'Hara, & P. Turner (Eds.), *Proceedings of the 31st international BCS human computer interaction conference (HCI 2017)* (pp. 1-5). ScienceOpen. https://doi.org/10.14236/ewic/HCI2017.18

Urbanek, M., & Güldenpfennig, F. (2019). Unpacking the audio game experience: Lessons learned from game veterans. In J. Arnedo, L. E. Nacke, V. Vanden Abeele, & P. O. Toups Dugas (Eds.), *Proceedings of the annual symposium on computer-human interaction in play* (pp. 253-264). ACM. https://doi.org/10.1145/3311350.3347182

Urbanek, M., Güldenpfennig, F., & Schrempf, T. M. (2018b). Building a community of audio game designers – Towards an online audio game editor. In I. Koskinen, & Y. Lim (Eds.), *DIS'18 Companion: Proceedings of the 2018 ACM conference companion publication on designing interactive systems* (pp. 171-175). ACM. https://doi.org/10.1145/3197391.3205431

Urbanek, M., Habiger, M., & Güldenpfennig, F. (2019). Drag'n'Hear: Creating, playing, and understanding audio games online. In J. Arnedo, & L. E. Nacke (Eds.), *CHI PLAY'19 Extended abstracts: Extended abstracts of the annual symposium on computer-human interaction in play companion* (pp. 125-130). ACM. https://doi.org/10.1145/3341215.3356986

Wang, R., Jung, C., & Kim, Y. (2022). Seeing through sounds: Mapping auditory dimensions to data and charts for people with visual impairments. *Computer Graphics Forum*, *41*(3), 71-83. https://doi.org/10.1111/cgf.14523

Welcome at AudioGames.net. (n.d.). AudioGames.net. Retrieved June 7, 2024, from https://audiogames.net/

White, G. R., Fitzpatrick, G., & McAllister, G. (2008). Toward accessible 3D virtual environments for the blind and visually impaired. In S. Tsekeridou, & A. D. Cheok (Eds.), *DIMEA '08: Proceedings of the 3rd international conference on digital interactive media in entertainment and arts* (pp. 134-141). ACM. https://doi.org/10.1145/1413634.1413663

Wilson, M. (2019, September 16). *Microsoft's trickiest product might be its most important*. Fast Company. https://www.fastcompany.com/90400552/microsofts-trickiest-product-might-be-its-most-important

Zdanowicz, G., & Bambrick, S. (2019). *The game audio strategy guide. A practical course*. Routledge. https://doi.org/10.4324/9781351016438