

Article

Speech Puzzles (Spuzzles): Engaging the Reduced, Causal, and Semantic Listening Modes for Puzzle Design in Audio Games

Emmanouel Rovithis , Agnes Papadopoulou , Vasileios Komianos , Varvara Garneli  and Andreas Floros 

Department of Audiovisual Arts, Ionian University, 49100 Corfu, Greece; a.papadop@ionio.gr (A.P.); vkomianos@ionio.gr (V.K.); vgarneli@ionio.gr (V.G.); floros@ionio.gr (A.F.)

* Correspondence: emrovithis@ionio.gr

Abstract: This paper proposes a novel approach to audio game design by introducing the concept of speech puzzles (spuzzles) to describe the utilisation of recorded voice for the creation of audio puzzles in ways that challenge players' different listening modes. In the fields of audio games and audio-interactive applications, speech serves instructive, descriptive, narrative, and in some cases—in the form of hints or quizzes—gameplay purposes by addressing users through language. The suggested approach of spuzzles extends this potential by including, besides encoded meaning, the acoustic properties of sound, thus engaging the user's causal and reduced listening modes in parallel with the semantic listening mode. An audio game consisting of four inherently different spuzzles was designed as proof of concept and tested by seven third-year students of Audiovisual Arts, who elaborated on their experience through a focus group semi-structured discussion. Despite their difficulty, the spuzzles were well accepted by most of the participants (5/7), whereas all participants agreed on their acoustic richness, need for concentration, and independence from pre-existing musical knowledge. Therefore, the authors suggest that the proposed design approach could serve as a paradigm for future research in the design of complex audio-based game mechanics.

Keywords: audio games; puzzles; audio puzzles; game design; puzzle design; listening modes



Citation: Rovithis, E.; Papadopoulou, A.; Komianos, V.; Garneli, V.; Floros, A. Speech Puzzles (Spuzzles): Engaging the Reduced, Causal, and Semantic Listening Modes for Puzzle Design in Audio Games. *Appl. Sci.* **2024**, *14*, 3858. <https://doi.org/10.3390/app14093858>

Academic Editor: Gianluigi Ferrari

Received: 7 March 2024

Revised: 27 April 2024

Accepted: 29 April 2024

Published: 30 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Audio games (AGs) are electronic interactive applications, in which sound is used as the main modality to represent information [1]. Their evolution differs significantly from that of video games, in which ocularcentrism has to a great extent limited sound to a merely ornamental role. AGs have managed to exploit the benefits of ever-changing new media, from text-to-speech computer software to handheld electronic devices, and from arcade consoles to mobile applications and augmented reality environments, thus shaping novel game mechanics and in some cases producing popular titles [2]. Yet, AGs are still a long way from addressing a wide audience as their visual counterparts do. Considering the documented benefits of audio interaction not only in the field of entertainment, but also in education, cultural heritage, art, and creativity, there is a great need to raise the general public's awareness about the AG genre [3].

For this purpose, AG designers should advance their understanding of the AG design process, in order to create challenging and exciting audio-interactive experiences. However, the AG designer community is rather small and has important challenges to overcome [4]. Urbanek and Güldenpfennig suggest that the use of sound constitutes a “change of modality” that fundamentally differentiates AG from video game prototyping, making it inherently difficult to draw upon knowledge and practices from the field of video game design, while the existing AG design guidelines are lacking scientific documentation [5]. Another challenge pertains to one of the main drives behind AG design, namely their accessibility for visually impaired people. On one hand, AGs that address

blind users have been criticised as being too simple [6]. On the other hand, by adding complexity to the audio representation of the game's facets, it becomes too difficult for sighted players who have not developed their listening skills sufficiently to accomplish the game's tasks without visual support [7]. Making the game not just accessible but truly inclusive, i.e., an experience as similar as possible for sighted and visually impaired players, requires a careful design of the sonification strategy from the start of the development process [8]. Last, the design of additional layers of audio information to create interactive immersive environments must fulfil the needs of both functionality and aesthetics [9]. In other words, the acoustic experience must not only serve the game's mechanics but also be artistically appealing.

This research aims to contribute to AG design research by investigating a new design approach for AG mechanics proposed by the authors, which employs the acoustic properties of speech for shaping puzzles that challenge players' different listening modes. As proof of concept, the game "Break the Audio Code" was designed to demonstrate different facets of this "spuzzles" (speech + puzzles) approach, from using unprocessed speech recordings to complementing them with background sounds, and from applying simple editing actions to heavily manipulating the recordings via audio effects, in order to engage players' semantic, causal, and reduced listening. The acceptance of spuzzles was evaluated in a playtesting session and a consequent semi-structured discussion by a focus group comprised of seven students of the Audiovisual Arts Department of the Ionian University in Greece. The authors are aware of the limitations of their research, since the number of the participants is small, their profile similar, and no quantitative data were analysed. In this research, a qualitative research approach was employed aiming to provide preliminary insights regarding the perception of spuzzles that could serve as the basis for their future systematic design and utilisation. Since spuzzles employ different listening modes and thus exercise players' listening skills, they could be used to address the aforementioned challenge that sighted players are failing to cope with complex AG mechanics (Figure 1).

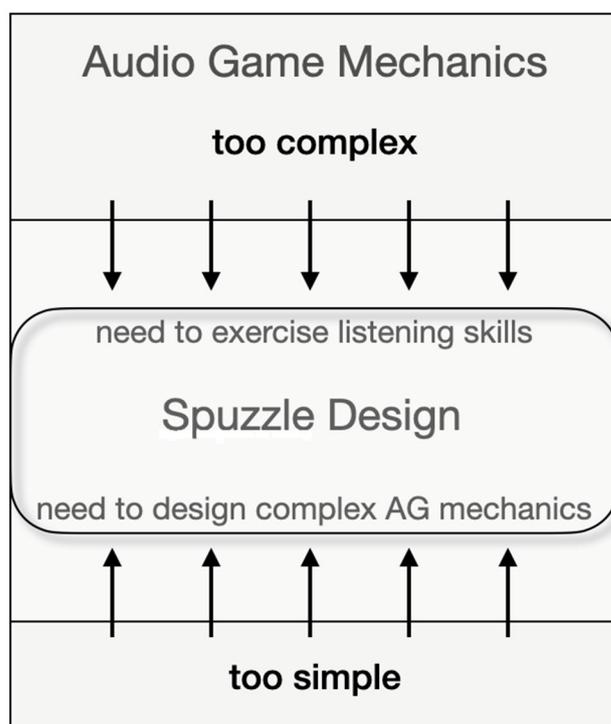


Figure 1. Research scope.

The rest of the paper is structured as follows. Section 2 presents the theoretical framework of this research focusing on its two main axes, (a) puzzles and (b) speech, in AG

design. Section 3 describes the game itself, as well as some salient design aspects. Section 4 analyses the authors' methodology for the evaluation of the experiment, and Section 5 presents the results in terms of the authors' observations, the participants' performance, and the content of the focus group discussion. Section 6 discusses the results, and Section 7 reports the preliminary conclusions. All data pertaining to this research, including game content and discussion recording, have been uploaded onto Open Science Framework (OSF) and are available via this link: https://osf.io/tcwea/?view_only=cb2824ad31284e05ac04f9889f50094e, accessed on 28 April 2024.

2. Background

2.1. Puzzles in AG Design

Discussing AG mechanics, Archambault et al. discern four different AG sub-genres: (i) action AGs, which test a player's timing and dexterity; (ii) exploration AGs, which involve on the navigation of auditory spaces; (iii) management/simulation AGs, which are based on the organisation of audio resources; and (iv) puzzle AGs, which present players with challenges that test their problem-solving skills [10]. In a broader context, game design experts Rollings and Adams define puzzles in general as "primarily mental challenges" [11]. The principles of audio puzzles are quite similar to those of video puzzles: solving them requires concentration, logic, strategy, pattern recognition, and other cognitive functions. Sometimes they are included in video games to intrigue players with their alternative gameplay [12]. In the context of AGs, audio puzzles can play a complementary role as necessary challenges for the plot to progress or a new space to unlock, or they can even constitute the game's main mechanic. In all cases, players need to listen carefully to the acoustic and/or musical properties of the audio stimuli, including pitch, timbre, dynamics, rhythmicity, consonance, and/or micro- and macro-structure, in order to find the solution.

Recorded and played back raw or processed through audio effects, speech lies at the heart of the puzzles concept. In AG puzzle design, the conventional utilisation of speech is limited to the inclusion of narrative elements (hints, questions, enigmas, etc.), which require players to comprehend the meaning of the spoken words. This mechanic challenges only the semantic listening mode. However, according to Chion, there are at least three listening modes: (a) semantic listening, when one interprets a message carried by a code or a language; (b) causal listening, when one tries to identify the sound's source and gather information about its generative process; and (c) reduced listening, when one focuses on the characteristics of the sound itself, ignoring its cause and meaning [13]. All three listening modes operate in tandem, are subject to subjective as well as objective criteria, and can be exercised and developed in the context of instructive experiences [14].

The proposed puzzle approach is novel in AG puzzle design because it extends the potential of speech beyond simple meaning to the acoustic properties and the realistic and symbolic references they can carry. Besides employing their semantic way of listening, players are thus required to employ their reduced and causal listening modes. This could enhance the educational and entertaining aspects of the genre. Liljedahl and Papworth suggest that with intelligent AG design, sighted players have the potential to overcome their difficulties in audio-only interaction [15]. This research aims to explore the meaningful inclusion of puzzles in AG design by investigating the acceptance of puzzles in an AG experience. As proof of concept, the authors designed an AG consisting of four puzzles, each with a different design approach, tested it with seven sighted participants, and extracted preliminary conclusions through a semi-structured discussion.

2.2. Speech in AG Design

Speech occupies just a portion of our everyday acoustic environment, yet it constitutes the primary medium for information distribution [16]. Research in the field of psychoacoustics has demonstrated the increased sensitivity of human hearing in the mid-frequency range (from lower-mid to higher-mid) [17], hinting at the species' evolutionary focus on the voice and the messages it expresses. In audio-visual media, the voice has been said to have

a captivating power compared to other elements, a phenomenon which has been called “vococentrism” [18]. When included in an audio-only environment, such as an audio film, words and phrases become so vital for understanding the action that they can distract from other non-verbal sounds or even overshadow them [19].

In AG design, speech is regarded as one of the three basic auditory elements along with music and sound, namely the one responsible for the transmission of knowledge: from the announcement of news or advice to the description of scenes, actions, and plot [20,21]. Researchers consider the dependency on verbal information to be one of the genre’s decisive factors, with some even discerning between two types of AG: those using spoken descriptions and those using only non-verbal audio cues [10,22]. Speech has also served as one of the main modalities to facilitate accessibility for the visually impaired. Since the early history of AG around the end of 1970s, any text-based game can become accessible to blind players via a screen reader [8].

The existing uses of speech in AG design include: (i) instruction, (ii) description, (iii) narration, and (iv) mechanics. These categories are not to be seen as disconnected from one another, but rather as intertwined manifestations of one common, inherent characteristic: they all rely on the semantic function of speech, on the meaning it carries. The following examples, drawn not only from AGs but also from the broader field of audio-interactive applications, demonstrate the broad range of the semantic use of speech.

On the most basic level, verbal audio delivers the necessary instructions for setting up and starting the game. Listening to the available options is particularly useful to visually impaired players who cannot rely on any graphical interface [23]. The instructions on how to enter the game are often complemented by a brief description of the goal and relevant tasks [24]. In other cases, an audio tutorial guides new players throughout their first contact with the game environment [25]. Spoken instructions are also included in the main game to describe the scene and remind players of current tasks [26]. This support can be in the form of an audio guide that introduces players to the important elements of each stage [27]. Balan et al. have reviewed various educational AGs on subjects including mathematics, programming, and biology, in which speech delivers instructions, announces the tasks at hand, and explains the educational content [28].

In terms of description, verbal audio is regarded as the most direct way to describe visual information. Albeit tedious and not without disadvantages compared to other sonification techniques, speech is more flexible in attributing subtle differentiations and/or sub-categories of an information structure [29]. Personal audio guides typically utilise spoken descriptions of cultural content to offer augmented-reality-enhanced tourist experiences [30]. Verbal description of points of interest can not only refer to the real world but also represent in detail a fictional world, such as the medieval setting with mages and monsters in [31].

The more an audio guide is combined with the game element of exploration, the more mere description of information is enriched with narrative elements. Players are told the game’s story and provided with cues as to the necessary route or other hints about required actions, to follow that story [32]. In that context, speech lies at the intersection of describing the game world and being part of the game’s mechanics. Players need to interpret the meaning of the spoken words, in order to find the right path and pursue the game’s objectives. An interactive audio story with both historical and fantasy elements can thus be shaped via monologues and/or dialogues, which are often binaurally recorded to enhance player’s immersion in the audio-represented world [2,33].

Another means for speech-facilitated mechanics is a spoken quiz, in which players are asked a question and must input the correct answer to enter the next stage [34]. Verbal output from the system to the user seems to be so well accepted that the input of voice commands has been suggested to facilitate two-way verbal interaction [35]. However, the need for sophisticated techniques to facilitate natural speech communication has been stressed [36]. Nowadays, text-to-speech (TTS) technology has been greatly advanced by

artificial intelligence that employs natural language processing (NLP) technologies to mimic the human voice and provide meaningful feedback [37].

In consideration of the importance of verbal audio, the authors of this research have investigated an alternative use for speech in the context of AGs, namely its use as a raw material for the design of audio puzzles. They suggest that the implementation of puzzles can shape new possibilities in AG experiences, since players are required to concentrate on the acoustic properties of speech, rather than only on its meaning. Speech has an intuitive impact on human psychology; it features a universal intimacy. It can thus serve as an efficient medium for training sighted players for complex audio-only interaction and potentially help them overcome their difficulty in completing challenging AG tasks.

The potential of puzzles to exercise players' listening skills is further enhanced by the educational function of games in general and puzzles in particular. The motivational power of gamification practices is widely accepted as their main benefit to the learning process [38]. However, research has recently started to focus on investigating the correlation between specific game mechanisms and learning outcomes [39]. Aligned with learning theories, such as problem solving and flow, findings have shown that challenging puzzles, even without the support of immersive elements (narratives, avatars), have a positive effect on learning both directly and via increasing engagement [40]. As a result, puzzle games have been applied to promote critical and creative thinking on a variety of topics [41–43].

To the authors' knowledge, there has been only one approach to exploring the potential of the acoustic properties of verbal information rather than its meaning. Walker et al. have suggested "spearcons", a technique of speeding up a spoken phrase until it loses its speech recognizability, as a flexible way to sonically express menu items [44]. However, the puzzles proposed in this paper differentiate themselves in various fundamental ways: (a) they pertain to the creation of audio puzzles rather than audio menus; (b) they can potentially employ all audio processing techniques rather than only variable playback speed; and (c) they do not eliminate the recognizability of speech, but instead retain the semantic connection with the represented information, in order to serve potential informative and narrative purposes in parallel.

3. Game Design

This section analyses the design of the four puzzles of the AG "Break the Audio Code" designed by the authors for the purpose of this research. The content of the game was sought in the typography collection of the Ionian University Museum in Corfu, Greece. The collection is situated at the Ionian University Department of Archives, Library Science, and Museology, and exhibits machines, tools, and objects related to various phases of the production of printed or typed archival material. More specifically, the exhibits fall under the categories: (i) typesetting, (ii) printing, (iii) bookbinding, and (iv) typing. One exhibit was selected from each of those categories, and a descriptive text was provided by the museum for each of the selected exhibits. The texts were narrated and recorded by the corresponding author (male voice), and these recordings served as raw material for the creation of four respective puzzles (Figure 2).

There were three major specifications in the design of the puzzles:

1. In terms of mechanics, they had to engage players' different listening modes: semantic, causal, and reduced.
2. They also had to highlight important aspects of the represented exhibit. This meant that conceptual connections had to be established between selected information and its sonic representation.
3. The audio outcome had to be clear to the listeners, providing all the necessary clues to solve the puzzle.

In consideration of the above, all recordings were first processed through "noise reduction" and "compression" effects, which removed any unwanted background ambience and increased the volume of the remaining voice. For the design of the audio puzzle, the methodology previously proposed by the authors was followed, according to which a

thorough research and analysis of the information to be sonified leads to designing the game's genre, levels, and mechanics [45]. The authors relied on two axes of description for each of the exhibit: (a) the item's purpose and historical importance in terms of the general category that it belongs to (typing, bookbinding, printing, typesetting), and (b) the item's specific information in the narrated text.

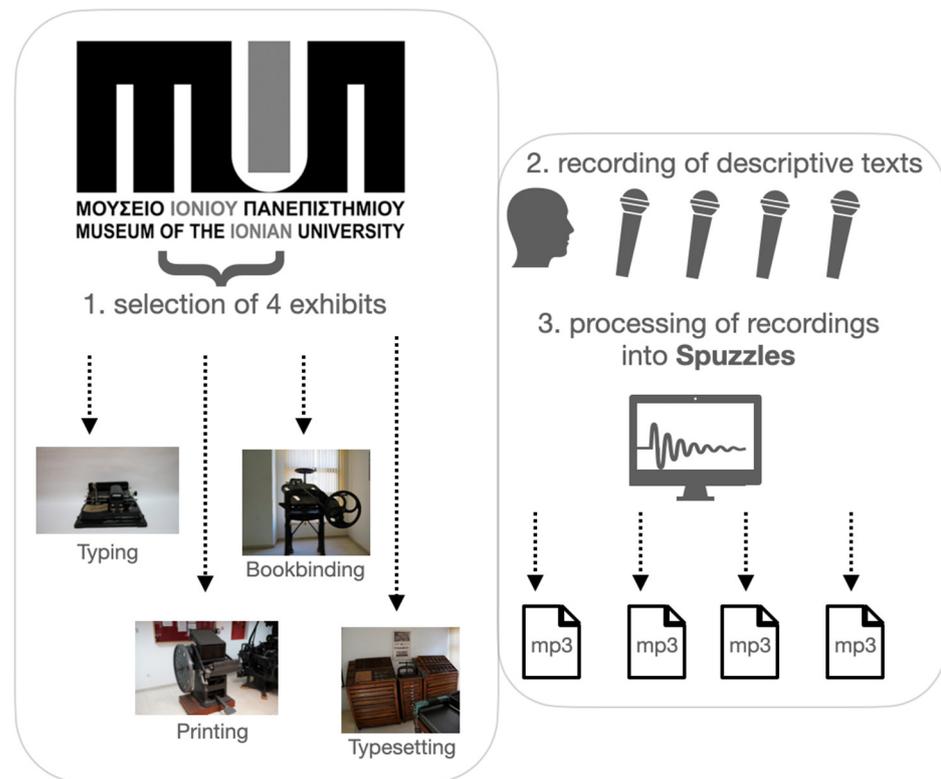


Figure 2. Game content.

It was decided that the solution of each puzzle be a two-digit number (10–99), as it is an easily manageable piece of information. Establishing semantic connections between sound and exhibit was quite simple since the source of information consisted of spoken language. The causal and reduced connections proved to be more complicated to forge, as elements of the item itself and the function it serves essentially had to be symbolically represented by audio properties and processes. Any audio effects used for this purpose should be unambiguous and avoid causing cognitive overload. From this perspective, the authors felt that their role as designers were reminiscent of that of a neurosurgeon, who performs a delicate operation.

The design of each spuzzle is analysed below:

Spuzzle 01. “Typesetting”—targeted listening mode: **Semantic**

The exhibit that was selected for this puzzle is a large piece of furniture with typographic cases with the following text narrated and recorded:

Within a printing house, each typeface had a family of letters (font), in all even sizes (from 6 to 72). The smaller drawers contained from the middle upwards Latin characters and from the middle downwards their respective punctuation marks.

The authors' intention was to highlight the main function of this exhibit, which is to store the letters and all other symbols for the typesetting process. Thus, the authors attempted to draw players' attention to one basic property of these characters, namely their size. The semantic approach was selected for turning these words into a puzzle. Players need to listen carefully and think about how many different font sizes are stored in the exhibit. Since only even sizes are mentioned, including both 6 and 72, the “solution formula” is $(72 - 6)/2 + 1$, which results in 34.

Spuzzle 02. “Typing”—targeted listening mode: **Causal**

The selected exhibit is an AEG Mignon 3 Typewriter, and it was accompanied by the following recording:

The AEG machine was manufactured in the year 1915 in Germany, and it is the Mignon 3 model. Its keyboard covers German-speaking countries (Latin characters) and works with a letter indicator. It is black, medium size, with a white keyboard, and weighs about 8.2 kg.

The typewriter is a device of the modern world. People are familiar with the object itself, more so with the act of typing, whether it is on an old-fashioned machine or a computer keyboard. The authors wanted to highlight the importance of this invention by means of the characteristic “clickety-clack” sound, so the sound of twenty clickety-clacks was placed in the background of the narration, spread out irregularly throughout its duration. Players need to discern the sounds behind the speech, employ their causal listening to understand that the sound source pertains to the exhibit, and count them.

Spuzzle 03. “Printing”—targeted listening mode: **Reduced**

The selected exhibit is a foot-operated electric printing press. The recorded text was processed as follows:

The foot-o’erated and electrically driven u’right ‘rinting ‘ress has an ink, the form moves and goes u’ and down, and it could be worked with the foot or with a motor that turned the wheel, and the ‘age was ‘rinted. Then, the clean ‘a’er was ‘laced for the next ‘rint.

Printing is also a process that many people relate with in their everyday life. The authors’ intention was to highlight the need for something printed to be complete and clear to the detail. When printing text, if even a letter is problematic, then the outcome is not satisfactory. This was acoustically represented by omitting a letter from the recording through editing. Players need to concentrate on the acoustic pauses and realize that they always occur on the letter ‘p’, which is number 16 in the Greek alphabet.

Spuzzle 04. “Bookbinding”—targeted listening mode: **Reduced**

The selected exhibit is the binding cutting machine “No 23959 Karl Krause Leipzig”. It was presented through the following text recording: *The Karl Krause Leipzig cutting machine No. 23959 was manufactured in the late 19th century and was a useful tool for the bookbinder. This cutter is suitable for cutting large quantities of paper in the same dimension. It also serves in cutting the three sides of the book equally when it was necessary.*

How a book is bound is something unknown to the majority of people. Thus, the authors wanted to highlight the process itself, as it is described in the text. For that purpose, the “chorus” audio processing effect was applied to most of the recording, up until “equally when it is necessary”. The chorus effect creates slightly out-of-tune copies of the initial sound. The authors created two such copies, one played back slightly faster and the other slightly slower, and placed them to the maximum left and right positions of the stereo panorama, respectively, while keeping the initial voice in the centre. As a result, players listen to three distinct voices, and when the effect stops only one voice remains. The point where this happens (bold letters) was carefully selected to express a metaphorical connection to the bookbinding process. The three voices played back through different speeds represent the three sides of the book, and when these are cut “equally” only one voice remains. The solution is from “3 to 1” thus 31.

As demonstrated in Table 1, two spuzzles were designed to target players’ semantic and causal modes of listening, respectively, the first one by presenting the recorded text unmodified and the second one by mixing it with a non-verbal background sound. The other two spuzzles target players’ reduced listening mode, one by muting parts of the recording, and the other by mixing the recording with two copies of itself played back at different speeds and panned to the extreme left and right sides of the acoustic space.

Table 1. Spuzzles’ sound design and scope.

Spuzzle	Applied Audio Process	Targeted Listening Mode
“Typesetting”	None	Semantic
“Typing”	Mixing (sound added in the background)	Causal
“Printing”	Silencing (specific audio parts muted)	Reduced
“Bookbinding”	Chorus & Panning (out-of-tune copies placed Left & Right)	Reduced

4. Evaluation Methodology

For conducting their investigation, the authors selected the qualitative research approach. Qualitative research aims at investigating in depth rather than breadth the experience of participants in natural, non-experimental settings [46]. The purpose of this research can be summed up in the following research question:

RQ: How do sighted players perceive spuzzles as an audio-only puzzle mechanic?

In qualitative research, the purposeful rather than random selection of a focus group composed of individuals who share connections with the phenomenon under examination maximises the understanding of their perspective [47]. By means of data collection techniques including observations and interviews, researchers attempt to dive into participants’ thoughts and thus gain important insight into the complexity of their shared experience. Thus, for this research, the selected methods included a focus group, observation of their performance, and semi-structured discussion about their experience.

The participants were recruited from the student body of the Ionian University Department of Audio and Visual Arts, where the authors teach. More specifically, a total of seven individuals (four males, three females, aged 20–23), all students of the fifth semester course “Interactive Multimedia”, were gathered in the classroom equipped with personal computers where the weekly course takes place. The participants had already attended a 3 h lecture on audio-interactive applications and audio games the week before.

In the beginning, the participants were informed about the terms and scope of the research: they would first fill in a technographic questionnaire, then split into teams and play a collaborative audio game of puzzles, and then take part in a semi-structured discussion about their experience. The discussion would be horizontal, meaning that they would be able to reply to the questions posed by the moderator, as well as comment on each other’s replies, openly and in no particular order. They were reassured that they can quit the process at any time and were asked to provide their consent in written form. No explanation was given prior to the game regarding its content or the spuzzle design method to ensure that the participants would not form any preconceptions.

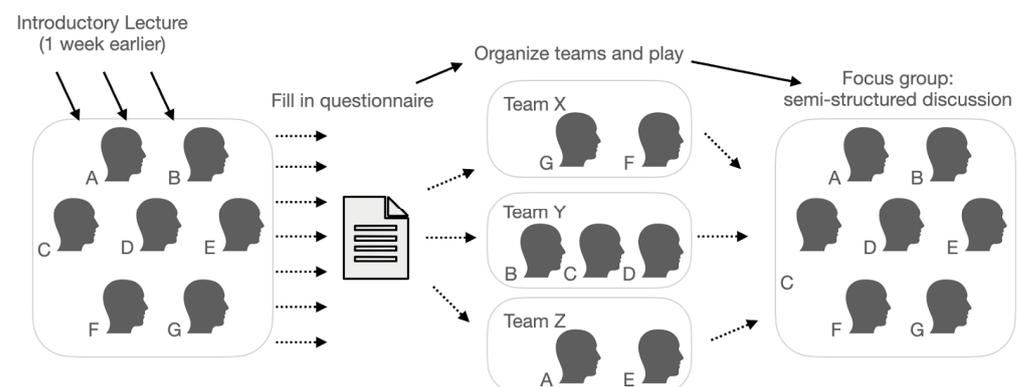
The technographic questionnaire helped researchers to expand their understanding of the participants’ background and thus orient them in an appropriate way towards the research goal. More specifically, the participants were required to estimate their own experience with (i) puzzle games (PGs), (ii) audio games (AGs), (iii) audio technology (AT), and (iv) music education (ME), using a scale from 1 to 4 (1 = none, 2 = a little, 3 = adequate, 4 = a lot). Table 2 demonstrates the participants’ profiles with the individual “scores” and their total sum. The two participants with the highest total score (F and G) are both very confident in the AT and ME categories. From the remaining participants, E and A demonstrate, respectively, the highest and lowest experience in AT and ME, while their self-evaluation in categories PG and AG is the same. Lastly, participants B, C, and D share similar profiles, feeling slightly more comfortable with the gaming aspect (PG and AG) than with that of sound (AT and ME).

Table 2. Participants' technographic profiles.

Participant #	XP in Puzzle Games	XP in Audio Games	XP in Audio Tech	XP in Music Education	Total
A	3	2	1	1	7
B	2	3	2	3	10
C	4	3	2	2	11
D	4	3	2	2	11
E	3	2	3	3	11
F	2	2	4	4	12
G	4	3	4	4	15

It must be noted here that after filling in the technographic questionnaire, participant B had to unexpectedly leave for a while. Based on the questionnaire findings, the authors divided the six remaining participants into the three following teams and decided that if participant B returned, he would join team Y (Figure 3):

- Team X included the “strong in sound” participants F and G;
- Team Y included the participants who were identical in all aspects and slightly more confident in gaming, C and D;
- Team Z included the participants A and E.

**Figure 3.** Organisation of the participants.

After shaping the teams, the authors explained the game rules to the participants. Each team would be stationed in front of a computer with four pdf documents on its desktop and be given four links to corresponding mp3 audio files, which have been uploaded to an online folder. The documents are locked and can only be opened if the correct two-digit password is entered. Each document includes the picture of the corresponding exhibit. Players need to listen to the mp3s and solve the puzzles to extract the two-digit passwords. There was also a fifth locked document in the author's laptop that would open if all four passwords were inserted in the correct order. That would be the chronological appearance of the typographical processes: typesetting–printing–bookbinding–typing. The team that first solves all four puzzles and opens the fifth document is the winner. The game would have a maximum duration of 1 h. Every 10 min, the teams get one chance to open each of the documents. This window of opportunity is a restriction that prevents participants from massively trying out different passwords without careful consideration. It would help them to organise their thoughts and prioritise potential solutions.

Regarding the fifth document, the inclusion of a meta-puzzle as the endgame is a very common technique in puzzle organisation [48]. Players are presented at once with several subsets of the meta-puzzle as prerequisites for the final solution and can solve them in any order they wish (Figure 4). The authors implemented this strategy in structuring the

game for two important reasons. First, a linear path for solving the puzzles would pose the danger that the latter would not be examined if the former were not solved. A non-linear meta-puzzle gives players motivation and freedom to plan their own strategy, whereas the opening of the in-between pdfs essentially reports on their progress. Second, to solve the meta-puzzle, players need to consider the chronological order of the exhibits, and in doing so they engage with their cultural context. This is aligned with a crucial factor proposed in the literature for enhancing learning in the field of virtual heritage, namely to establish a relationship between users, virtual content, and cultural context [49].

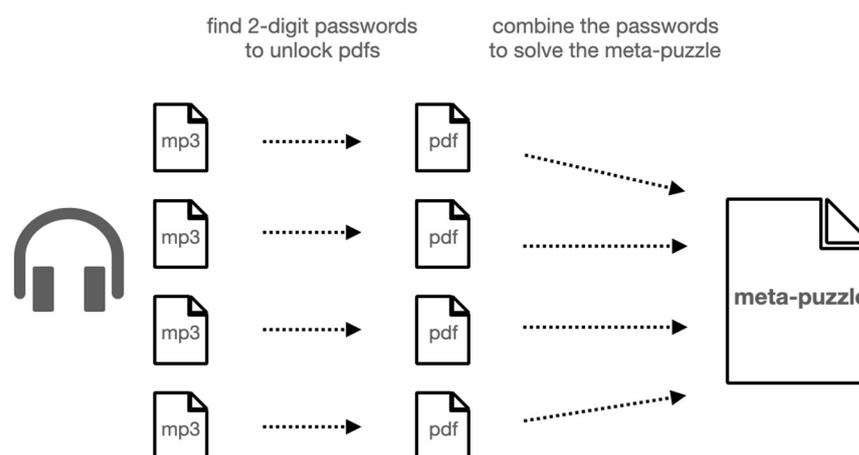


Figure 4. Game's non-linear structure.

After providing the procedural information, the authors provided the narrative context. The participants would play the role of activists that oppose a corrupted tycoon, who has illegally amassed a typographical collection. They have broken into his office and are looking for the password to hack into his cryptocurrency server and destroy his financial empire. They have planted a glitch in the alarm system that allows them to enter potential passwords every 10 min. The authors implemented this fictional setting to excite the participants' curiosity and fantasy and enhance their immersion in the game process. Storytelling is a fundamental element of game design [50] that has also gained a lot of attention in the context of cultural visits [51].

All of the above steps were completed in 40 min, after which the game started. During the game, the author made notes of his observations. The winning team completed the game in 40 min (fourth opportunity window). After a small break of 15 min, the session continued with the discussion, which was recorded for further analysis.

The discussion was semi-structured with open questions. The authors had predefined some criteria to analyse the participants' replies. Thus, they would be able to focus on specific factors, while being open to any new ones that might come up. These criteria were:

- RQ-cr01.: the participants' background;
- RQ-cr02.: the participants' enjoyment;
- RQ-cr03.: the game's perceived difficulty;
- RQ-cr04.: the game's perceived potential to exercise listening skills.

Throughout the discussion, the participants were encouraged to comment on each other's thoughts and ideas. All solutions to the spuzzles were also explained and discussed. Occasionally, the author would address the more "quiet" participants to ensure that their opinions were heard as well. In general, the discussion revolved around the following questions:

- What did you think of the experience?
- Would you play a game of spuzzles again?
- What do you think about the game's difficulty?
- What did you like and/or not like?

- What would you change in the design of the game?
- What kind of strategy did you follow (in general, and/or in each specific spuzzle)?
- What did you think about the voice as the main carrier of information?
- What do you think about the game enhancing your connection to the exhibit?
- What do you think about the game developing your acoustic perception?

5. Results

This section presents the results of the experimental playtest session, which include the authors' observations noted down throughout the process, the participants' performance in the game, and their feedback drawn from the recorded focus group discussion.

5.1. Researcher's Observations

At the beginning of the game process, all teams were particularly silent and seemed focused. Each team had only one set of headphones, so team members listened to the puzzles in rotation. Very often they remained contemplative before sharing their reasoning with the other team members. In each 10 min round, multiple listenings of the puzzles were performed. All teams were taking notes of the possible solutions and were quiet when talking to each other so as not to annoy the other teams.

As the game progressed, teams X and Z seemed to become more joyful, as they were talking louder, smiling, and sometimes laughing. Both these teams were eager to try out new passwords and asked for more frequent opportunity windows. On the other hand, team Y remained mostly silent with face expressions and body language sometimes showing frustration. Around 25 min into the game, team Z asked for some help and all teams enthusiastically agreed. In response, the author suggested that all teams are granted one clue for a puzzle of their choice. Then, he thought of and noted down the following clues to provide if asked:

- Bookbinding = "Does it sound like what it does?" (a hint at the metaphor between the audio process and the typographical function);
- Typing = "Is an author really alone?" (a hint at the existence of the typing sound in the background);
- Printing = "Print it again, please!" (a hint at the faulty printing result due to the missing letter);
- Typesetting = "We 've got all that fits your needs!" (a hint at the described variety of different fonts).

At the end of the game, teams X and Z burst into applause, whereas team Y seemed relieved. They all seemed a little exhausted from the effort, therefore the author said they could have a 15 min break.

5.2. Participants' Performance

Team X was the winner by completing the game in the fourth round, in around 40 min game time. In their first attempt, they solved the typing and bookbinding puzzles, and in the next round the typesetting spuzzle. After three failed attempts with the printing spuzzle, they asked for a clue and thus managed to solve it in their fourth attempt.

The author allowed the other teams to complete their fourth attempt as well. Team Z had already solved two spuzzles: typing in the first and printing in the second round. Then, they asked for the typesetting clue but did not manage to find the correct passwords in the remaining rounds. Team Y had solved the typing spuzzle in round two and then asked for the printing clue. In the last attempt, they managed to solve the bookbinding spuzzle. In the meantime, after round 1, participant B had returned and joined team Y, which was at that time behind in score. Table 3 demonstrates all teams' progress.

Table 3. Teams' progress.

Spuzzles Solved in:	Round 1	Round 2	Round 3	Round 4	Sum of Spuzzles Solved
Team X	Typing Bookbinding	Typesetting	-	Printing	4 (winner)
Team Y	-	Typing	-	Bookbinding	2
Team Z	Typing	Printing	-	-	2

5.3. Semi-Structured Discussion

In general, the participants were positive towards the game. "It was interesting", said participant A. "A nice experience. I liked that we had to pay attention not only to the technical characteristics of sound, the pauses, the audio processing, but also to the meaning of the words", commented F. The other participants nodded in agreement about this dual aspect of the game. "I also liked the backstory that we were spies, it provided us with a nice context", said E, who further suggested: "I would enhance the backstory and the game in general through visual information, and maybe make it into an electronic application". Participant F added: "If you were to make it into a game application, many things should be visualised through text or animation, there could be a specially designed interface to listen to the files and insert the passwords and then unlock the final puzzle", with B suggesting that "maybe even have a narrator telling the backstory like in some mobile apps".

The discussion continued about features that in the participants' opinion would embellish or improve the game. "It would be nice to have a visual countdown or a limited number of attempts", suggested A. "You would then be motivated by more pressure", added E. "There could also be a reward system, for example fewer attempts means more stars", said G. "Or you could win more time and more hints", commented participants A and E. Participant C suggested "a loud feedback about right and wrong actions to make the game more intense", and G mentioned "better recording quality if you were to release this game commercially".

Most of the participants (5/7) agreed that they would play an audio game with spuzzles again. "Yes", said participants A, C, and E; "definitely", said F; "of course", said D. Participant B, though, was negative regarding the idea and said, "I wouldn't play a game of puzzles, because I am not good at them, it doesn't have to do with audio", whereas G "would play again a game of puzzles, but not audio ones". The participants were asked to elaborate. "I also had a weak internet connection", said B. "The first puzzle was easy, but then the game became increasingly difficult and one had to retain complex information in their mind", commented A. Her teammate E added, "the first puzzle we tried was easy. When we listened to the rest we thought that there were different levels of difficulty and we got lost".

Then, each spuzzle was played back, and the participants were asked to elaborate on the strategy they followed to solve them. Regarding the "typing" spuzzle, F explained, "besides the voice there was this distracting sound of the typewriter, so I counted the clickety-clack". His teammate G "stopped paying attention to the voice, as soon as the first clack was heard". "We solved it the same way, it was in fact the first puzzle I listened to", commented A, whereas her teammate E had "first listened to all puzzles and noticed that in one puzzle there is the clickety-clack, in another the sudden pauses, so when I listened again I suspected it didn't have to do with the words, but rather with their background. What confused me was that some clacks were heard from the left and others from the right, so maybe it had to do with panning". "What can also confuse is that each clickety-clack sound is double, one press and one release sound", commented F, and C added, "the two clacks were too close to each other". "Coming from the speakers, the sounds were clearer than from my headphones", said G, but B disagreed. "I could tell that the double

sound counts as one event, since I have had experience with a typewriter toy as a child”, mentioned F.

Regarding the “bookbinding” spuzzle, E said, “this was a nightmare. We tried to count the seconds of the audio effect’s duration, we also counted the amount of processed words”. “We definitely noticed there was a change at some point”, added her teammate A, “but we would never think to combine two one-digit numbers (3 and 1) into one two-digit number”. “We solved it in a technical way, 3 voices into 1”, said F. “We solved it in a different way, the words said that three sides are cut into one book”, commented C.

Regarding the “typesetting” spuzzle, F explained, “we didn’t notice anything technical, so we kept notes of the numbers that can be inferred from the meaning”. B complained, “this and all other puzzles required fantasy and I am not good at it”. E said, “I was on the right track, but made a mistake in the calculation”.

In the “printing” spuzzle, participant D “couldn’t find any clues unlike in the other puzzles”. “At first we counted the pauses” said F, and A agreed, saying, “we then noticed that the pauses were always silencing the same letter”. “We found the clue we asked for very useful”, added G.

In terms of difficulty, most of the participants agreed that the game was hard but not too hard. “At first I was stressed, because I thought that the game would need musical knowledge, but then I realised that it had to do with the perception of non-music sounds”, said A, adding that the “difficulty had more to do with making logical connections, to which sound was actually helpful”. Participant C agreed that “musical knowledge was neither required nor developed, all one needed was their acoustic perception”. “And attention”, added F, “you really need to focus”. “Yes, your full concentration was required”, agreed C.

Asked about the connection mediated by the game between player and exhibit, participant B did not agree, suggesting that “this kind of puzzles could be designed for any artifact or text whatsoever”. “It provided with some knowledge, but I think most of us focused only on the necessary information to solve the puzzle”, said C. Participant A disagreed, arguing that “there was a connection, for example in “Typing” there was the sound of the typewriter, so you could assume that the rest of the puzzles build on similar relationships to the exhibits”. Her teammate E added that “it wasn’t just a puzzle. There was both an educational aspect of sound that can be applied to any subject, and specific connections to the exhibits. These connections would be stronger if the game took place in the physical exhibition space”.

Lastly, regarding the impact of human voice as the main carrier of game information, participant F said that “it made the game more familiar compared to using a melodic or abstract sound. In addition, if we had just to count the clickety-clack, it wouldn’t be so interesting, there wouldn’t be a dialogue between sound sources of different nature”. Participant A commented, “it was stressful to listen to the same voice over and over again, but in the end it made the game better, you had to solve it to make the voice stop”. “Maybe next time add more voices to choose from”, suggested E. “Nevertheless, I would remove the three-voice effect, it drove me mad”, commented B, with G arguing for “less extreme stereo panning”, and A disagreeing since “this was the best part”.

Figure 5 presents a short overview of the participants’ feedback.

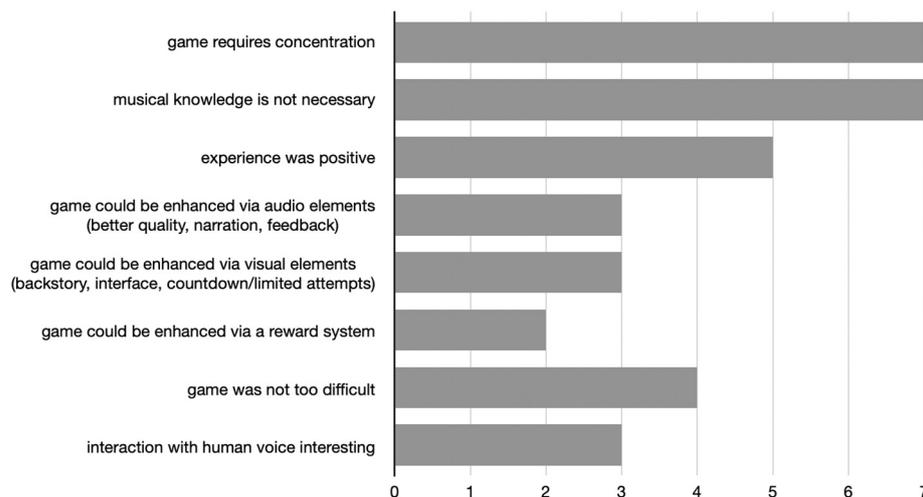


Figure 5. Overview of participants’ feedback.

6. Discussion

Most participants (5/7) enjoyed the game and were positive towards a similar experience in the future. From the two participants that did not quite share this opinion, one excluded the audio factor from his reasoning and admitted a general dislike of puzzles. The other, even though he did not provide any explanation when asked, was critical throughout the discussion towards the quality of the game’s audio recordings and not of the audio puzzle mechanic (Figure 6). His critical attitude may be connected with a more general stance exhibited by the whole group regarding high expectations of electronic games. However, most suggestions for game improvement were not related to audio-based gameplay but to video game conventions, such as countdowns, rewards, and narratives, more so in their visual form. This finding implies the impact of modern video game culture, as well as the ocularcentrism of the electronic game industry. It also indicates that players are not familiar with the act of playing only by means of the acoustic modality.

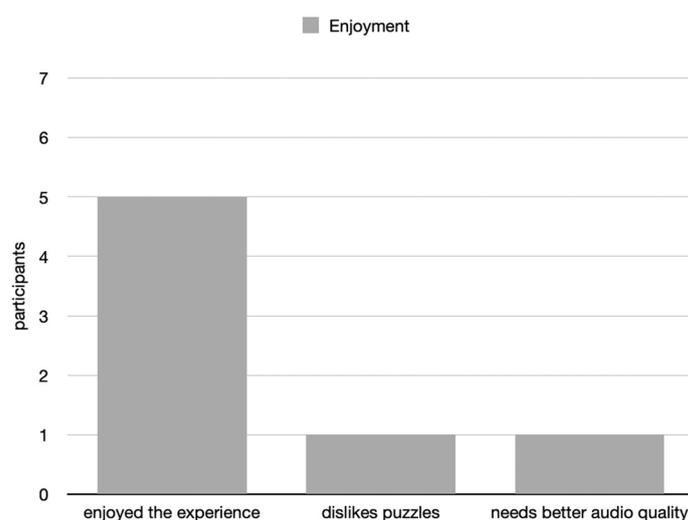


Figure 6. Results regarding players’ enjoyment.

There were two things that all participants agreed on: (1) the speech puzzles require players to fully concentrate on the acoustic properties of incoming audio, and (2) this challenge can be addressed by everyone regardless of their musical experience.

The state of concentration has been frequently connected to AG experiences, not only in terms of a required condition for game progress but also as a skill that players

can develop through the process [22]. Focusing on sound's learning affordances, Bishop, Amankwatia, and Cates, among others, report increased attention and the exclusion of distractions, yet they argue that, however frequent, sound's use in learning environments deals with the literal conveyance of information, while neglecting the associative potential that would facilitate a deeper study of the learning material [52]. The participants in this research seemed to comprehend the multifaceted dimension of the game's audio content. They recognised and sought for other aspects of the audio content besides the meaning of speech, such as background stimuli, disturbances, pauses, repetitions, and stereo positioning. Afterwards, they discussed and criticised the perceived sound properties in detail. Moreover, they seemed appreciative of the fact that the game's mechanics relied solely upon their acoustic perception and not their musical knowledge. This instilled a feeling of justice in the experience, i.e., all players were equal, since all that was needed was to employ one's everyday listening skills. However, the authors see two sides in that condition. On the one hand, the exclusion of musical requirements can indeed provide a more open framework in terms of player address. On the other hand, it cannot be ignored that the winning team was comprised of the two participants who felt the most confident in their music and audio technology skills.

The implied assumption that a stronger background in the music and audio technology fields results in better performance in audio puzzles needs to be thoroughly investigated in the future. However, it must also be noted that none of the teams lost their focus on the game goal. In fact, even Team Y, which had shown signs of frustration, was persistent enough to solve their second puzzle just before the end. Participants' commitment in the game process despite its increased difficulty may be accounted for by the familiarity of the human voice. Participants felt comfortable with, and in a stimulating way competitive against, the notion of a talking person. This is aligned with findings in the literature about the positive disposition of audio game players towards different in-game voices, because they felt like they were getting to know real people [26].

In terms of the employed listening modes (Figure 7), the fact that all teams solved the "typing" spuzzle quite early is indicative of the importance of causal listening. All participants were able to distinguish between two different sound sources, trace their causality, and make connections to the exhibit. However, one could argue that the perception of the clickety-clacks in the background does not relate to the acoustic properties of speech per se. The spuzzles best suited to answering this would be the ones employing the reduced listening mode. Judging from their performance, the sighted players of this research do not seem less capable of coping with reduced listening than with semantic listening challenges. In fact, only one team solved the semantic listening spuzzle "typesetting", whereas two teams solved the reduced listening spuzzles "printing" and "bookbinding". Regarding the latter, Team X solved it by employing just their reduced listening skills, whereas Team Y relied mostly on their semantic listening skills. This demonstrates the flexibility and creative freedom of the spuzzle design approach, which can target multiple listening modes in parallel.

Combined with recording, editing, and processing techniques, human speech gains an acoustic richness that listeners can engage with. All participants, even if they had not found the solution to a spuzzle, could identify and discuss the nuances of its acoustic parameters and possible ways to interpret them. This acoustic richness can provide a creative arsenal for AG designers to shape complex audio interactions based on different listening modes. A multitude of acoustic and musical aspects, like timbre, pitch, amplitude, rhythmicity, etc., as well as a multitude of audio-processing effects, such as chorus, delay, pitch-shift, distortion, etc., can be utilised towards meaningful gameplay and to address players' acoustic perception. Therefore, the authors suggest that the creative use of recorded human voice for puzzle design holds great potential to foster the exercising and development of players' listening skills.

Another issue that arose pertains to listeners' experiential connection to the audio content. The most ambiguous sound proved to be the clickety-clack of the typewriter, with

some participants having difficulty in understanding the shape of one single event. The participant who relied on his personal experience with a similar sound source successfully identified the sound's two-event morphology. This implies that players who are experientially related to the sounds of an audio puzzle could have an advantage. To mitigate such imbalances, it is the sound designer's responsibility to shape a carefully designed audible outcome with clear literal and/or metaphorical references. To achieve this, the real nature of sound sources and sound events may have to be in some cases enhanced or augmented, and in others questioned, distorted, or even redefined.

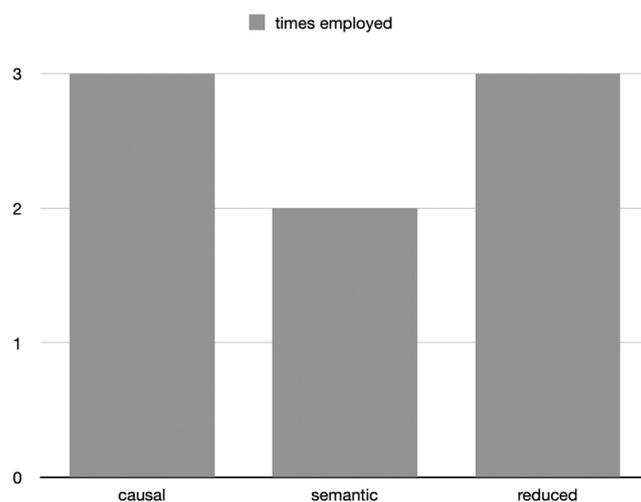


Figure 7. Successful employment of different listening modes.

7. Conclusions

In this paper, the authors have presented the puzzles AG design approach to demonstrate that the use of human speech in AG design is not restricted to the aspect of language but can also include the acoustic aspect to shape complex audio-only interactions. Whether in the form of raw recordings or processed through audio effects, speech can serve as the material for audio puzzle gameplay that challenges players' causal and reduced listening modes, while retaining its semantic function, which is traditionally responsible for delivering the game's instruction, description, and narration. Thus, acoustic or even musical concepts can be literally or metaphorically connected to the sonified data.

The proposed design approach was investigated through a playtest session and a semi-structured discussion with a focus group of seven students of audio-visual arts in the context of an interactive multimedia course. The results have shown that the puzzles were well accepted, their acoustic richness perceived, the imposed need for concentration identified, and their independence from musical knowledge appreciated. They were mostly criticised because of their audio-only form, which does not match the visually oriented standards of the video game industry.

The authors are aware of the limitations of their research, which can only provide preliminary insights regarding the acceptance of this new AG design method. The participants agreed that the puzzles required their full concentration and that they could participate regardless of their pre-existing musical knowledge. Combined with the existing literature on the inherent features of audio interaction, these findings are a promising indication of the puzzles' potential to exercise sighted players' listening skills, and thus help them cope with complex AG tasks. For a deeper investigation of this method's educational efficiency, as well as its potential to address sighted and visually impaired players alike, more targeted and systematic research must be conducted in the future. However, the authors hope that the prototype presented here will inspire AG designers and grant them with new creative possibilities regarding the realisation of exciting audio-interactive experiences.

Author Contributions: Conceptualization, E.R.; methodology, E.R., A.P. and V.G.; software, E.R.; investigation, E.R.; resources, E.R. and V.K.; data curation, E.R.; writing—original draft preparation, E.R.; writing—review and editing, A.P., V.K. and V.G.; supervision, A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki. Ethical review and approval were waived for this study. The participants were students of the “Interactive Multimedia” course, in the context of which (as well as other courses in the Department) they have frequently participated in similar user experience evaluations.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://osf.io/tcwea/?view_only=cb2824ad31284e05ac04f9889f50094e accessed on 28 April 2024.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Parker, J.R.; Heerema, J. Audio interaction in computer mediated games. *Int. J. Comput. Games Technol.* **2008**, *2008*, 178923. [CrossRef]
- Beksa, J.; Fizek, S.; Carter, P. Audio games: Investigation of the potential through prototype development. In *A Multimodal End-2-End Approach to Accessible Computing*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 211–224.
- Rovithis, E.; Floros, A.; Moustakas, N.; Vogklis, K.; Kotsira, L. Bridging Audio and Augmented Reality towards a new Generation of Serious Audio-only Games. *Electron. J. E-Learn.* **2019**, *17*, 144–156. [CrossRef]
- Urbanek, M.; Guldenpfennig, F.; Schrempf, M.T. Building a Community of Audio Game Designers—Towards an Online Audio Game Editor. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, MD, USA, 20 October–1 November 2017; pp. 171–175.
- Urbanek, M.; Guldenpfennig, F. Rethinking prototyping for audio games: On different modalities in the prototyping process. In Proceedings of the 31st British Computer Society Human Computer Interaction Conference, Sunderland, UK, 3–6 July 2017; p. 18.
- Giannakopoulos, G.; Tatlas, N.-A.; Giannakopoulos, V.; Floros, A.; Katsoulis, P. Accessible electronic games for blind children and young people. *Br. J. Educ. Technol.* **2018**, *49*, 608–619. [CrossRef]
- Östblad, P.A.; Engström, H.; Brusck, J.; Backlund, P.; Wilhelmsson, U. Inclusive game design: Audio interface in a graphical adventure game. In Proceedings of the 9th Audio Mostly: A Conference on Interaction with Sound, Coimbra, Portugal, 7–9 September 2011; pp. 1–8.
- Engström, H.; Brusck, J.; Östblad, P.-A. Including Visually Impaired Players in a Graphical Adventure Game—A Study of Immersion. *IADIS Int. J. Comput. Sci. Inf. Syst.* **2015**, *10*, 95–112.
- Friberg, J.; Gårdenfors, D. Audio games: New perspectives on game audio. In Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Singapore, 3–5 June 2005; pp. 148–154.
- Archambault, D.; Ossmann, R.; Gaudy, T.; Miesenberger, K. Computer games and visually impaired people. *Upgrade* **2007**, *8*, 43–53.
- Rollings, A.; Adams, E. *Andrew Rollings and Ernest Adams on Game Design*; New Riders: Indianapolis, IN, USA, 2003.
- Pichlmair, M.; Kayali, F. Levels of sound: On the principles of interactivity in music video games. In Proceedings of the DiGRA Conference, Tokyo, Japan, 24–28 September 2007.
- Chion, M. *Audio-Vision: Sound on Screen*; Columbia University Press: New York, NY, USA, 2019.
- Chion, M. *The Three Listening Modes. The Sound Studies Reader*; Columbia University Press: New York, NY, USA, 2012; pp. 48–53.
- Liljedahl, M.; Papworth, N. Beowulf field test paper. In *Audio Mostly*; ACM: Piteå, Sweden, 2008; p. 43.
- Drossos, K.; Floros, A.; Kanellopoulos, N.-G. Affective acoustic ecology: Towards emotionally enhanced sound events. In Proceedings of the 7th Audio Mostly Conference: A Conference on Interaction with Sound, Corfu, Greece, 26–28 September 2012; pp. 109–116.
- Suzuki, Y.; Takeshima, H. Equal-loudness-level contours for pure tones. *J. Acoust. Soc. Am.* **2004**, *116*, 918–933. [CrossRef]
- Chion, M. *The Voice in Cinema*; Columbia University Press: New York, NY, USA, 1999.
- Lopez, M.J.; Pauletto, S. The Sound Machine: A Study in Storytelling through Sound Design. In Proceedings of the Audio Mostly Conference, Piteå, Sweden, 15–17 September 2010; pp. 1–8.
- Roden, T.E.; Parberry, I.; Ducrest, D. Toward mobile entertainment: A paradigm for narrative-based audio only games. *Sci. Comput. Program.* **2007**, *67*, 76–90. [CrossRef]
- Röber, N.; Masuch, M. Interacting with Sound: An Interaction Paradigm for Virtual Auditory Worlds. In Proceedings of the 10th Meeting of the International Conference on Auditory Display, Sydney, Australia, 6–9 July 2004.

22. Targett, S.; Fernström, M. Audio games: Fun for all? All for fun. In Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, 6–9 July 2003.
23. Moffat, D.C.; Carr, D. Using audio aids to augment games to be playable for blind people. In Proceedings of the Audio Mostly Conference-A Conference on Interaction with Sound, Piteå, Sweden, 22–23 October 2008; pp. 35–42.
24. Miller, D.; Parecki, A.; Douglas, S.A. Finger Dance: A Sound Game for Blind People. In Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, MD, USA, 20 October–1 November 2017; pp. 253–254.
25. Garcia, F.E.; de Almeida Neris, V.P. Design Guidelines for Audio Games. In *Human-Computer Interaction. Applications and Services: 15th International Conference, HCI International 2013, Las Vegas, NV, USA, 21–26 July 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 229–238.
26. Song, D.; Karimi, A.; Kim, P. Toward designing mobile games for visually challenged children. In Proceedings of the International Conference on E-Education, Entertainment and e-Management (ICEEE), Bali, Indonesia, 27–29 December 2011; pp. 234–238.
27. Nagele, A.N.; Bauer, V.; Healey, P.G.; Reiss, J.D.; Cooke, H.; Cowlshaw, T.; Baume, C.; Pike, C. Interactive Audio Augmented Reality in Participatory Performance. *Front. Virtual Real.* **2021**, *1*, 610320. [[CrossRef](#)]
28. Balan, O.; Moldoveanu, A.; Moldoveanu, F.; Dascalu, M.-I. Audio games—a novel approach towards effective learning in the case of visually-impaired people. In *ICERI2014 Proceedings*; IATED: Valencia, Spain, 2014; pp. 6542–6548.
29. Blum, J.R.; Bouchard, M.; Cooperstock, J.R. What’s around me? Spatialized audio augmented reality for blind users with a smartphone. In Proceedings of the Mobile and Ubiquitous Systems: Computing, Networking, and Services: 8th International ICST Conference, MobiQuitous, Lisbon, Portugal, 20–25 November 2011; pp. 49–62.
30. D’Auria, D.; di Mauro, D.; Calandra, D.M.; Cutugno, F. A 3D audio augmented reality system for a cultural heritage management and fruition. *J. Digit. Inf. Manag.* **2015**, *13*, 203–209.
31. Lyons, K.; Gandy, M.; Starner, T. *Guided by Voices: An Audio Augmented Reality System*; Georgia Institute of Technology: Atlanta, GA, USA, 2000.
32. Röber, N. Playing Audio-only Games: A compendium of interacting with virtual, auditory Worlds. In Proceedings of the 2005 Digra International Conference: Changing Views: Worlds in Play, Vancouver, BC, Canada, 16–20 June 2005.
33. Paterson, N.; Naliuka, K.; Jensen, S.K.; Carrigy, T.; Haahr, M.; Conway, F. Design, Implementation and Evaluation of Audio for a Location Aware Augmented Reality Game. In Proceedings of the 3rd International Conference on Fun and Games, Leuven, Belgium, 15–17 September 2010; pp. 149–156.
34. Fakhour, M.; Azough, A.; Kaghat, F.Z.; Meknassi, M. A cultural scavenger hunt serious game based on audio augmented reality. In *Advanced Intelligent Systems for Sustainable Development (AI2SD’2019) Volume 1-Advanced Intelligent Systems for Education and Intelligent Learning System*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 1–8.
35. Boletsis, C.; Chasanidou, D. Smart Tourism in Cities: Exploring Urban Destinations with Audio Augmented Reality. In Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference, Corfu, Greece, 26–29 June 2018; pp. 515–521.
36. Härmä, A.; Jakka, J.; Tikander, M.; Karjalainen, M.; Lokki, T.; Hiipakka, J.; Lorho, G. Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc.* **2004**, *52*, 618–639.
37. Tsepapadakis, M.; Gavalas, D. Are you talking to me? An Audio Augmented Reality Conversational Guide for Cultural Heritage. *Pervasive Mob. Comput.* **2023**, *92*, 101797. [[CrossRef](#)]
38. Kapp, K.M. *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training And Education*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
39. Dicheva, D.; Dichev, C.; Agre, G.; Angelova, G. Gamification in education: A systematic mapping study. *J. Educ. Technol. Soc.* **2015**, *18*, 75–88.
40. Hamari, J.; Shernoff, D.J.; Rowe, E.; Coller, B.; Asbell-Clarke, J.; Edwards, T. Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Comput. Hum. Behav.* **2016**, *54*, 170–179. [[CrossRef](#)]
41. Padilla, J.J.; Lynch, C.J.; Diallo, S.Y.; Gore, R.J.; Barraco, A.; Kavak, H.; Jenkins, B. Using simulation games for teaching and learning discrete-event simulation. In Proceedings of the 2016 Winter Simulation Conference (WSC), Washington, DC, USA, 11–14 December 2016; pp. 3375–3384.
42. Gintere, I. A new digital art game: The art of the future. In *Society. Integration. Education. Proceedings of the International Scientific Conference Vol. 4*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 346–360.
43. Adams, V.; Burger, S.; Crawford, K.; Setter, R. Can you escape? Creating an escape room to facilitate active learning. *J. Nurses Prof. Dev.* **2018**, *34*, E1–E5. [[CrossRef](#)] [[PubMed](#)]
44. Walker, B.N.; Lindsay, J.; Nance, A.; Nakano, Y.; Palladino, D.K.; Dinger, T.; Jeon, M. Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Hum. Factors* **2013**, *55*, 157–182. [[CrossRef](#)] [[PubMed](#)]
45. Rovithis, E.; Floros, A.; Mniestris, A.; Grigoriou, N. Audio games as educational tools: Design principles and examples. In Proceedings of the 2014 IEEE Games Media Entertainment, Toronto, ON, Canada, 22–24 October 2014; pp. 1–8.
46. Neuman, D. Qualitative research in educational communications and technology: A brief introduction to principles and procedures. *J. Comput. High. Educ.* **2014**, *26*, 69–86. [[CrossRef](#)]
47. Merriam, S.B. Introduction to Qualitative Research. *Qual. Res. Pract. Ex. Discuss. Anal.* **2002**, *1*, 1–17.
48. Nicholson, S. Peeking Behind the Locked Door: A Survey of Escape Room Facilities. In White Paper. 2015. Available online: <http://scottnicholson.com/pubs/erfacwhite.pdf> (accessed on 20 April 2024).

49. Bekele, M.K.; Champion, E. A Comparison of Immersive Realities and Interaction Methods: Cultural Learning in Virtual Heritage. *Front. Robot. AI* **2019**, *6*, 91. [[CrossRef](#)]
50. Jenkins, H. Game design as narrative architecture. *Computer* **2004**, *44*, 118–130.
51. Chatzidimitris, T.; Kavakli, E.; Economou, M.; Gavalas, D. Mobile Augmented Reality edutainment applications for cultural institutions. In Proceedings of the IISA 2013, Piraeus, Greece, 10–12 July 2013; pp. 1–4.
52. Bishop, M.J.; Amankwatia, T.B.; Cates, W.M. Sound's use in instructional software to enhance learning: A theory-to-practice content analysis. *Educ. Technol. Res. Dev.* **2008**, *56*, 467–486. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.