Going beyond the Visual Domain

Improving the UX of Sonic Interaction

Sohrab Gharibpour
Abstract

Sonic Interaction is an area of Interaction Design which has the focus on the sound related interactive system designs. This area of interaction design has not been researched as much as visual interactive systems and because of this, sonic interactive systems are not as much developed and enhanced; This study tries to examine the possibility of improving and enhancing the sonic interactive systems by combining Embodied interaction with Sonic interaction. The main concern of this study is to combine the user’s head movements in a manner of embodiment with sonic interactive systems and explore the feasibility of improvement and enhancement of sonic interactive systems in combination with embodied interaction. For this reason, data from three experiments has been gathered and evaluated into the conclusions of this study.

Keywords: Sonic, Interactive, System, Design, Embodied, Spatial, Sound, Head, Gesture, Game

1. Introduction

Interacting with Sonic systems has a long story and goes back to the first days of the automobile engineering, when engineers had to listen carefully to the sounds of the engine and compare it to the sound of other engines, analyzing the rhythm for discrepancies (to determine and fix potential issues) (Franinović & Serafin, 2013). Even today, drivers regularly listen to the sounds of the engine and interact with automobile through sound.

Sound has different aspects which define how the sound is shaped and what impression it has. Regarding music, as an example of more specific form of sound, there are 7 aspects which are called “Pitch”, “Dynamics”, “Rhythm”, “Articulation”, “Timbre”, “Order” and “Tempo” (Marin & Deutsch, 1982). These are the artistic aspects of the sound which are being used by musicians and sound designers to define the impression of the sound. But there is more about the sound than these very familiar aspects which are less explored. Moving away from the artistic view of sound and getting closer to everyday life sounds that we deal with every day, we face these features and aspects of sound which, as HCI and Interaction designers, we should consider knowing them and using them to deliver a better user experience to the end users of interactive systems and specially sonic interactive systems.

Every sound has a source which is located in the 3D environment, meaning it has a 3D coordination. In response to this feature we have a sense called “Spatial Hearing” (Blauert, 1985) which helps us locate the sound source using our ears. A sound might spread in all directions or it might be focused in specific direction. The velocity of the sound also varies, depending on the listener’s and sound source’s movement. The change of a sound’s velocity results in an effect called “Doppler effect” (Chowning, 1971) which changes the sound’s impression on the listener.

Another parameter which strongly affects the sound and its impression is the “Reverberation” effect (Paterson et al., 2011). “Reverberation” effect explains how the shape
and frequency of sound waves change depending on the size, material and other properties of
the environment where sound is being played. This is the reason most recording studios have
isolated sound padding to reduce reverberation, while adding a reverberation effect to a sound
can modify the listener’s impression of the sound’s environment.

All the features and properties mentioned above explores the 3D nature of sound. This is
where the exploration of a realistic Sonic Interaction, which resembles the everyday life
experience for the user, begins.

3D sound is a very ubiquitous and embodied phenomenon which could be best described
and explored using “Embodied Interaction” (Dourish, 2001). It is a theory and framework
which explores and describes how “Embodied Interactive Systems” should be designed and
investigated. On the other hand “Sonic Interactive Design” (Rocchesso & Serafin, 2009) is an
area of “Interaction Design” and “Human-Computer Interaction” fields which explores the
design of sound-related (Sonic) interactive systems.

The main concern of this research paper is the exploration and understanding of realistic
sonic interaction which I call “Embodied Sonic Interaction” and that’s the area where “Sonic
Interaction Design” and “Embodied Interaction Design” meet. Regarding this, as the general
question through this paper, we are trying to answer the question “How can Embodied
Interaction be used to Improve and Enhance Sonic Interaction?”

1.1 Embodied Interaction

*Embodied Interaction* (Dourish, 2001) is a philosophical point of view on interaction design
and human-computer interaction. According to this theory designers should consider the
embodied nature of human interaction when designing and creating interactive systems
(Dourish, 2001).

Embodied interaction focuses on the nature of human interaction with his environment in
opposition to focusing on the technology (Dourish, 2001). In other words, embodied
interactive systems should be designed and created in the way that users feel natural while
interacting with the system.

One of the notions of embodiment is tangible computation (Dourish, 2001, p. 189). Tangible
computation tends to move the computation from the traditional desktop computer stations
into the environment to make it possible for the users to interact with these interactive systems
in a natural way that they do in the world and on the move (Dourish, 2001, p. 189). To be
clearer, while interacting with a tangible computation system, users use their senses and
embodiment and receive the response and feedback from the tangible computation system.
Wearable devices could be named as an example of tangible computation systems.

The main concern of current study is to examine the possibility of integration of sonic
interaction with embodied interaction in context of tangible computation to improve and
enhance the user experience of sonic interaction. In this paper we have studied the human’s
head movements as an embodied way of interaction and combined theme with sonic
interactive systems to deliver a tangible design and explored three contexts of sonic interaction
which could be improved or enhanced to deliver a better interaction user experience. These
three contexts are, human-computer communication through the sonic interaction and
human’s head movements; translating user’s head movements into a virtual 3D environment for improving the artificial 3D sound’s experience; and finally, possibility of game design using embodied interaction in the form of head movements and sonic interaction in the form of virtual 3D sound environment.

2. Related research

Sound is one of the very important tools for interaction. Interaction designers and human computer interaction designers have always been arguing for the reasons why sound should be considered as a tool for interaction. “Clearly, a good first reason to use sound is simply because it’s there” (Gaver, 1989, p. 2).

Sound is a very strong tool and “the appropriate use of non-speech sounds has the potential to add a great deal to the functionality of computer interfaces.” (Gaver, 1989, p. 2).

Also, interactive systems are more and more moving towards wearables and body tracking instruments as a natural way of interaction. There are many new techniques and technologies being developed as the new ways for interacting and controlling of interactive systems designed by interaction designers and human computer interaction designers. Kinect and Leap Motion and other similar technologies are examples of instruments being used for a technique called whole body tracking. Whole body tracking is the process of tracking users body motions and using them as a way of interaction with interactive systems (England et al., 2009). Whole body tracking could also be considered as a way of embodied interaction (Dourish, 2001) when it comes to designing interactive systems and that’s the area of interest in this research paper to combine embodied interaction with sonic interactive systems.

Sound is everywhere and exists in every movement and body action. Every simple interaction that we do in our everyday life involves some sort of sound generation and sound interaction. Regarding these embodied nature of sound, it is a very potential phenomena to be explored and investigated with an embodied interaction point of view to get it more involved in future designs of embodied interactive systems.

Several studies and experiments have been done by researchers in the field of sonic interaction. Among these studies, some have explored and investigated the embodied nature of the sound in its most basic form while others have examined and experienced interactive 3D sounds.

Two case studies are presented here which we will investigate in the following sections. The first case study is an experimental research formed in a series of workshops, followed by exercises and group projects meant for studying the extensibility of Sonic Interaction Design in everyday objects, in an embodied approach. The second case study explores the topic of feasibility of 3D sound as a mobile application suitable for urban navigation.

2.1 Sound Embodied

In an experimental study held by Karmen Franinovic, Daniel Hug and Yon Visell in 2007, researchers designed a series of workshops and invited a wide range of international participants, including designers, students and researchers. As the results of these workshops
they wanted to investigate “new roles for auditory display in everyday products” (Franinovic, Hug, & Visell, 2007, p. 335), and also for finding possible methodologies for designing these everyday products. In the beginning, researchers described their assumptions of what part sound can play in product design and potential expectations for expanding the use of these products in the future (Franinovic et al., 2007). The assumed use of the products are listed as following:

Ways to create or reveal functionalities in a product:
- Displaying new informational capacities (relaying or translating information sonically).
- Revealing and relaying information previously unavailable (such as affordances in an augmented artifact).
- Relaying information about product functionality in a manner that relies on other senses than vision (Brewster, 2002, pp. 188–205).

Reshaping the “sonic appearance” of an artifact:
- Consider augmenting the emotional aspects of the design (Norman, 2003).
- Enhancing the interplay-relations of material, shape, size and actions.

Improving the performance and usability within the interaction process:
- Aid user control of interfaces (Williamson & Murray-Smith, 2005, pp. 45–52), tools, devices (Rath & Rocchesso, 2005, pp. 60–69) or physical activities by providing feedback (Effenberg, 2005, pp. 53–59).
- Sonify “silent information” (pulse, speed, involuntary reactions) associated with certain actions (eg. biofeedback).
- Considering and improving the experience flow and focus of a user during a task.

Based on the roles explained above, researchers defined some exercises and activities for the participants. These activities included lectures for the participants to receive initial knowledge to work with sonic systems, programs and accessories. For later exercises, participants were asked to do “Ear Cleansing” exercises, in which participants would hear different samples of sound, and later they were asked to recall and describe what they had heard. The exercise covered the important basic skills for any design activity, which according to the researchers is “the ability to use the senses in question, and to be able to abstract, communicate and conceptualize about the experiences related to them” (Franinovic et al., 2007, p. 336).

After the exercises, participants formed project groups. Every group was meant to do a short project. Project results were covering a wide variety of interactive sound concepts, from interactive sound maps meant for urban sonification to musical compositions, video scenarios and prototyping (Franinovic et al., 2007).

We will briefly explore the interactive sonic prototypes created through the workshop as samples of how interactive sonic ideas could be developed through embodied interaction.
2.1.1 Interactive prototype projects

The first project mentioned in the study was named “Game-all-over” and was created by Benjamin Janke. It involves transforming unreal virtual instruments into the real world stuff using tangible and sonic objects (Figure 1). An example of such an artifact is a jump-pad that mimics the idea of throwing the player into the sky (Franinovic et al., 2007) (Figure 2).

![Figure 1. “Game-all-over” presents the idea of transforming unreal virtual instruments into the real world stuff using tangible and sonic objects. (Franinovic et al., 2007, fig. 4)](image)

The second project was named “Sonic Lockers” and was created by Stefano Teseo and Barbara Schuler. The idea was to explore the idea of sonically augmenting lockers at a train station, changing the symbolism of locking, opening and closing the lockers. The lockers had sounds that made these interactions somewhat exaggerated, such as increasing the “weight” of the locking mechanism or making it sound like you’re dropping your suitcase down a fiery chute (Franinovic et al., 2007) (Figure 4).

“Klingenstrasse” aka “Ringing Street” referring to the name of a street in Zurich created by Song Vega and Luisa Beeli is the next project created as part of the experiment. The idea of Klingenstrasse was basically to augment the natural sounds on the street. The concept was based on an analysis of sonic and spatial properties of the street itself. The project worked around the idea of enhancing the sounds of the street and design further sounds that would activate when cars passed over it, or when strong winds were blowing. (Franinovic et al., 2007).

Daniel Senn, David Herzog and Johannes Kiesbauer created a project called “Thirsty bottle” which, presented with a video scenario, was supposed to be a sonically augmented bottle aimed at make it inviting to drink from it. It would use sounds of sparkling fluids and music associated
with the different beverages and brands poured into the bottle. It was supposed to enhance the user experience of the drinking beverages (Franinovic et al., 2007).

“Sonic Fishing” created by Marcel Tanner, Daniel Volzke and Daniel Fischer, used the idea of sonically enhancing an ordinary fishing pole. The concept was developed by using a wide array of methods, including body storming and an augmented video scenarios in which the pole would use fish finder sensors, able to detect sound underwater and information about

Figure 2. User trying the prototype called “Game-all-over”. A jump-pad that mimics the idea of throwing the player into the sky. (Franinovic et al., 2007, fig. 5)

Figure 3. “Sonic Fishing” uses the idea of sonically enhancing an ordinary fishing pole. (Franinovic et al., 2007, fig. 7)
quantity, size and distance of fish - as well as the hook and bait used by the fisherman (Franinovic et al., 2007) (Figure 3).

The last project explained as part of the study was called “Gamelunch” created by Stefano Fumagalli, Stefano delle Monache, Stefano Papetti and Simone Lüling. Sonic Dining was a project focusing on augmenting an ordinary dinner experience to make it more lively and interactive. The Sonic Dining table would feature equipment that allowed it to perform a responsive soundtrack during the meal (Figure 5). The everyday objects on the table would be given new meaning, as interaction with plates, cutlery and the table itself would result in feedback from the system in the form of different kinds of audial responses to encourage creative expression and a performative atmosphere (Franinovic et al., 2007) (Figures 5, 6).

### 2.1.2 Workshop results

The main three conclusions of the workshops described by workshop holders and researchers could be summarized as, interactive sound design (or better said, sonic interaction design) is a very time consuming process and needs significant time to be spent for experimental design. Also, sonic interaction is a very complex and wide area of research and, regarding this fact it, requires at least a reasonable understanding of the field. Also, in addition to basic knowledge for the designers of the experiments and researchers, it would be an advantage for the

![Figure 4. “Sonic Lockers” examines the idea of a sonically enhanced locker at a train station.](image-url)
participants to have a basic understanding of the area of research, as they need to try a wide variety of different tools and techniques during the experiments and activities (Franinovic et al., 2007).

Based on the workshops and group projects, workshop organizers and researchers in the field came up with the following conclusions.

**Figure 5. User trying the “Gamelunch” prototype.** (Franinovic et al., 2007, fig. 8)

**Figure 6. The main focus of “Gamelunch” prototype is on augmenting a usual dinner experience to make it more interactive.** (Franinovic et al., 2007, fig. 9)
2.2 Feasibility of spatial sound

In the second experiment, a group of researchers formed their study based on Spatial Audio Navigation (Hinge, Jensen, & Poulsen, 2012). In this experiment, researchers denoted navigation by walking towards a sound. They created a prototype in the form of an Android application which uses the device’s GPS sensor to locate the user on the map in combination with Google’s navigation API to track the user’s path and compare it to the shortest possible path. They also used the OpenAL library which is meant to be used for efficient rendering of multichannel three-dimensional positional audio (http://en.wikipedia.org/wiki/OpenAL). They used the API from this library for producing spatial sounds which guides the users towards the target location (Hinge et al., 2012).

The application designed for this research is a friend finder which guides two users from two distanced locations, using the application at the same time to find each other by walking toward the directions that application shows them by playing spatial sounds (Hinge et al., 2012).

The prototype was created and developed in 3 iterations consisting of one technical iteration in which researchers investigated the possibilities of implementing the prototype by exploring the Android NDK, Android SDK and OpenAL API. It also involved a video prototype to conceptualize and communicate the design idea and interaction form, and helped define the requirements for the project. For the second iteration, researchers invited 5 test users and let them go through the video prototype with a free surf approach (Wiberg, 2005). They also came up with new ideas for applicability of the resulting application and prototype. As the third iteration, they started implementing the actual prototype based on the results from the previous iterations (Hinge et al., 2012).

Before being able to evaluate the prototype, researchers had to consider several issues. They improved the spatial sounds for the OpenAL library’s Head Related Transfer Function (HRTF) (Zotkin, Hwang, Duraiswaini, & Davis, 2003) to clarify if the positioned sounds is in front of, or behind, the users. Users reported that it is clearly obvious when the sound is positioned to the right or left of the users, but when it was in front or behind them they had difficulty determining the sound’s point of origin (Hinge et al., 2012).

To improve the user experience (and as a hint toward the correct direction), researchers added a pitch changing mechanism to the system. Using this mechanism, the pitch of the spatial sound would change when the user was moving toward the relative correct or opposite direction. By listening to the pitch of the sounds, users would know for sure that they were moving in the correct direction. The algorithm used for this calculated the distance between two users and changed the pitch of the sound based on the distance reduction or increment.

The researchers evaluated the prototype with a group of 13 participants in a controlled environment. Another group of 18 participants evaluated the prototype in real world scenarios. Results indicated that the test subjects in all cases were able to find each other, and that the distance they walked exceeded the shortest route between the two starting points only on the average of 7% compared to the shortest route from Google navigation API. Results from the
evaluations showed that researchers achieved the precision of 17° and 25° and an accuracy of 3° and 10° in a controlled and a real world experiment respectively (Hinge et al., 2012).

3. Methodology

Considering previous studies, experimental research (Robert L. Solso, M. Kimberly Maclin, MacLin, & Solso, 2007) in the form of controlled experiments is one effective way of exploring and investigating the sonic interaction field. Controlled experiments are an appropriate method for accurate measurements of usability levels. Controlled experiments are more accurate than data collected during informal testing or field evaluation (Newman, Lamming, & Lamming, 1995). Taking this into account, for the results to be more valid and accurate, evaluations were designed with the theory of controlled experimentation in mind - effectively trying to hold the conditions constant during the whole experiment and for every participant (Robert L. Solso et al., 2007). The experiment also uses a wide background range of participants regarding country of origin, culture, academic and technical background to have a more varied group (Robert L. Solso et al., 2007). The experiment is designed to pursue the issue, “Does the chosen design strategy meet requirement R?” (Newman et al., 1995). Following the “Experiment Design” (Newman et al., 1995) patterns, the experiment is designed towards generating more accurate results.
4. Studies

In this section, the process of designing the experiments will be further explained, along with the results of the individual experiments. Later sections will provide a further analysis of the gathered data, followed by subsequent conclusions.

4.1 Design

While exploring the related studies done by other researchers, several ideas of embodied interaction in combination with sonic interaction were explored. Generally these ideas are not designed for a real life scenario and are mostly experiments showing the possibility of involving the sonic interaction in everyday life, but embodied interaction is a much stronger framework for designing technologies for everyday life and real user scenarios.

For the purpose of this study, the experiments were designed to consider three challenges which could improve usability and user experience of sonic interactive systems in combination with embodied interaction framework and is a step forward towards the everyday life embodied sonic interaction.

The three challenges which were explored through this study could be explained in the form of these three questions:

1. Could embodied interaction in the form of head movements be used to interact with a Sonic Interactive System?
2. Could embodied interaction in the form of head movements be used to improve Spatial Sonic Interactive Systems?
3. Could embodied interaction in the form of head movements and Sonic Interaction be used for developing new methods of gamification?

The main focus of these studies is the usability of user’s head movement as a very natural way of interaction and specifically, embodied interaction for humankind. The idea is that user’s head movements could be used in two general ways of interaction with the sonic interactive systems.

A pure sonic interactive system is assumed to be independent of any kind of visual interactive interface, head movements (in the form of head gestures) could be used as a form of interaction. Examples of these interactions could be confirmation or denial gestures in reply to the questions asked by the sonic interactive system, in the scenarios where user cannot answer the questions verbally.

The second method of interaction would be in an interactive system which tracks user’s head movements and generates spatial sound in response. To clarify, the human brain considers the binaural hearing (Xu, Li, & Salvendy, 2007) in combination with head orientation to guess the correct location of the sound source. The process of characterizing how an ear receives a sound from a point in 3D space is called HRTF (Zotkin et al., 2003). Scientists and computer engineers have developed computer algorithms to simulate these types of spatial sounds, but the problem with these algorithms is that (similar to the reports from the second case study) many other software developers have reported that these algorithms are just good enough for simulating the spatial sounds positioned on either sides of user’s head and not in front or behind. In other words, users can not differentiate between the spatial sounds
positioned in front or behind them. Trying to address this problem, the idea is to combine user’s head movements with this algorithm which is being used for generating spatial sounds. In other words, make the spatial sounds responsive to the user’s head movements, so the user’s brain would calculate the spatial sound’s location easier and more accurately.

To examine these theories and pursuing the research questions stated, three different experiments were designed. Every experiment started by explaining the process of the experiments to participants and the three experiments including interviews afterwards took between 45 to 60 minutes for every participant in total.

### 4.2 Yes/No/Maybe/Almost Sonic Questionnaire

The first experiment was intended to answer the question “How do people answer Yes, No and Maybe, only using head movements?” The experiment was designed to figure out what the natural way of answering these questions were for every participant. For this reason, users were asked to do a simple task. They were asked to wear a headphone with an IMU sensor mounted on top and listen to the questions through a pair of headphones. After listening to every question, they had to first answer the question by moving their head and then speak the answer verbally. Participants were aware that the IMU sensor recorded their head movements as they replied to the questions. The questions were designed in a way that could be answered as yes, no, maybe or almost. After the IMU sensor recorded the head movements in reply to every question the oral replies were also written down by the researcher. The results were analyzed and compared to initial assumptions and individually summed up into conclusions.

<table>
<thead>
<tr>
<th>Answer</th>
<th>Assumed Head Movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>![Head Movement Image]</td>
</tr>
<tr>
<td>No</td>
<td>![Head Movement Image]</td>
</tr>
<tr>
<td>Maybe/ Almost/ Some Times</td>
<td>![Head Movement Image]</td>
</tr>
</tbody>
</table>

Table 1. Initial assumptions about head movements in reply to questions.
The initial assumptions for these answers in regard to the head movements are shown in table number 3. These are the regular way of answering yes, no, maybe or almost replies by the researcher.

The experiment was done with 10 participants from 8 different countries and 8 different cultures. The results are shown in table number 4.

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Country of origin</th>
<th>Yes</th>
<th>No</th>
<th>Maybe/ Almost/ Some Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Female</td>
<td>Iran</td>
<td><img src="image" alt="Yes Head Movement" /></td>
<td><img src="image" alt="No Head Movement" /></td>
<td><img src="image" alt="Maybe/ Almost/ Some Times Head Movement" /></td>
</tr>
<tr>
<td>26</td>
<td>Female</td>
<td>Finland</td>
<td><img src="image" alt="Yes Head Movement" /></td>
<td><img src="image" alt="No Head Movement" /></td>
<td><img src="image" alt="Maybe/ Almost/ Some Times Head Movement" /></td>
</tr>
<tr>
<td>33</td>
<td>Female</td>
<td>Canada</td>
<td><img src="image" alt="Yes Head Movement" /></td>
<td><img src="image" alt="No Head Movement" /></td>
<td><img src="image" alt="Maybe/ Almost/ Some Times Head Movement" /></td>
</tr>
<tr>
<td>26</td>
<td>Male</td>
<td>Spain</td>
<td><img src="image" alt="Yes Head Movement" /></td>
<td><img src="image" alt="No Head Movement" /></td>
<td><img src="image" alt="Maybe/ Almost/ Some Times Head Movement" /></td>
</tr>
<tr>
<td>31</td>
<td>Male</td>
<td>Iran</td>
<td><img src="image" alt="Yes Head Movement" /></td>
<td><img src="image" alt="No Head Movement" /></td>
<td><img src="image" alt="Maybe/ Almost/ Some Times Head Movement" /></td>
</tr>
</tbody>
</table>
The results for this experiment (Table 4.) show that in a wide variety of cultures, head movements used to reply “Yes” and “No” to an answer are very commonly used as moving the head up and down repeatedly for answering “Yes”. For answering “No”, people usually move their head from side to side repeatedly. However, the answer to a question which is not clearly “Yes” or “No” (such as “Maybe”, “Almost” or something similar), people from different cultures use different head movements.
Therefore regarding the fact that computer programs, mostly ask users for confirmation in the form of “Yes/No” or “OK/Cancel” questions, these two types of questions could be answered using head movements which in general could be called “Head Gestures”.

![Image](image.png)

**Figure 8. Prototype created for evaluation using a cellphone mounted on top of a headphone**

### 4.3 Finding the Sound’s Source

Spatial Sounds are artificially generated sounds created by computer algorithms. These algorithms use a series of calculations called “Head-Related Transfer Function (HRTF)” (Xu et al., 2007) to simulate the 3D positioned sound sources in the space using the difference between the sounds heard through two ears using headphones. Assuming the user’s head is located at the zero point on all three axis in 3D coordination, and the y axis is pointing upward the user’s head, this algorithm manages to generate the 3D positioned sounds when the sound source is located somewhere on the x-z plan and not above or below this plane. This means that it is not truly possible to simulate sounds below or above user’s head in a way that they can locate the sound just by listening to them.

For solving this problem, we combine the HRTF algorithm with head movement tracking. For this purpose, a head mounted IMU sensor would detect the head’s current orientation and informs the HRTF algorithm to update the angle between the sound source and 3D coordination system’s axis. This gives the user the possibility to explore the virtual 3D space more accurately by rotating and tilting their head and find the exact location of the sound source.
To be able to evaluate this theory, we created a prototype which consisted of the IMU and the HRTF program. In this experiment, participants would hear 2 series of 3D positioned sound. The first series of these sounds were static, meaning that they would not get updated in response to user’s head movements and the second series of these 3D positioned sounds were dynamic, meaning that their position would get updated in response to user’s head movements and virtually simulate a more natural 3D sound experience.

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Static sounds position</th>
<th>Dynamic sounds position</th>
<th>Average response time: Statics</th>
<th>Average response time: Dynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Female</td>
<td>R U U L U R</td>
<td>R L F R B</td>
<td>0.54 Seconds</td>
<td>2.46 Seconds</td>
</tr>
<tr>
<td>26</td>
<td>Female</td>
<td>R U U L U R</td>
<td>U R L F R B</td>
<td>1.91 Seconds</td>
<td>7.92 Seconds</td>
</tr>
</tbody>
</table>

Figure 9. Participants trying to guess the sound’s location. (All the participants have consented to including their pictures in this paper)
Table 3. Results for Finding the Sound Source evaluations.
(F = Front, R = Right, L = Left, B = Back and U = Up)

Users were asked to listen to each positioned sound for 10 seconds and say where the sound is coming from as soon as possible both for static and dynamic sounds respectively. The responses from users were timed and these timing in addition to the directions mentioned by the users are collected in the table 3. The data is explained more in details in Data analysis and results section.

### 4.4 Trying a Sonic-only Game

For the third evaluation, a prototype was created through the results from two previous prototypes in the form of a game. The game works as a sound-only challenge combining the 3D sound and IMU head movement tracking into a survival form of game.

The games begins with an introduction welcoming the user to the sound only virtual reality game and then after a 3 seconds silence, a huge virtual lager gone position on a random location in the surrounding space would be heard warming up then shooting towards the user.

User’s task is to listen carefully for the direction where the laser beam is getting close to them, and try to face the laser beam as soon as they find the correct direction. The moment they faced the laser beam, they should tilt their head to either sides, as to avoid the shot.

If the user managed to do these steps in 3 seconds (with sufficient angular tilt) they would hear the laser beam passing. If performed incorrectly, the laser beam would hit the user and play the sound of an explosion.

It was almost impossible (for users who tried the game for the first time) to escape from the first laser beam, even though the game routine had been explained to them and they had already had some experience with head movement controlling the computer system. They had no clear mental model (Benyon, 2010) of the experience before being hit by the first laser beam. The only players who could avoid the first shot had already tried the game before.
The evaluation for the game prototype took place with all the participants from the two previous experiments. The time from the first laser hit to the first successful escape from laser beams (in addition to the number of escapes and explosions for a maximum 2 minutes gameplay) were recorded as shown below. (Actual gameplay time was also recorded in case someone didn’t want to keep playing)

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Time from first hit to first successful escape</th>
<th>Number of saves</th>
<th>Number of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>Female</td>
<td>5 Seconds</td>
<td>8 Saves</td>
<td>8 Hits</td>
</tr>
<tr>
<td>25</td>
<td>Female</td>
<td>5 Seconds</td>
<td>8 Saves</td>
<td>8 Hits</td>
</tr>
<tr>
<td>27</td>
<td>Male</td>
<td>0 Seconds</td>
<td>14 Saves</td>
<td>2 Hits</td>
</tr>
<tr>
<td>23</td>
<td>Female</td>
<td>5 Seconds</td>
<td>12 Saves</td>
<td>4 Hits</td>
</tr>
<tr>
<td>33</td>
<td>Female</td>
<td>10 Seconds</td>
<td>12 Saves</td>
<td>4 Hits</td>
</tr>
<tr>
<td>29</td>
<td>Female</td>
<td>5 Seconds</td>
<td>12 Saves</td>
<td>4 Hits</td>
</tr>
<tr>
<td>26</td>
<td>Male</td>
<td>40 Seconds</td>
<td>3 Saves</td>
<td>12 Hits</td>
</tr>
<tr>
<td>26</td>
<td>Female</td>
<td>10 Seconds</td>
<td>13 Saves</td>
<td>3 Hits</td>
</tr>
<tr>
<td>31</td>
<td>Female</td>
<td>0 Seconds</td>
<td>15 Saves</td>
<td>1 Hit</td>
</tr>
</tbody>
</table>

Table 4. Results for Trying a Sound-Only Game evaluations.

Figure 10. Participants playing the sonic game.
(All the participants have consented to including their pictures in this paper)
5. Data analysis and results

In the following sections I will analyze the data gathered through every experiment and investigate the outcomes of the data analyzed.

5.1 Results for Yes/No/Maybe/Almost Sonic Questionnaire experiment

Looking at the results from the first experiment, 100% of the participants used the same gesture for answering “Yes” questions. The gesture is performed by slightly moving the head down and up 2 to 3 times. The average amount of head reciprocations were 2 for more than 50% of the participants (Table 4).

Referring to the assumptions table (Table 3), only one participant used the first gesture (which is moving the head downward once) in reply to only one of the questions. Considering that 10 participant answered 15 questions each, 1 answer in general would be considered 1/150 which results in less than 1% occurrence. Because of the low probability, this data was excluded from the results and 100% is considered the usage of the second gesture suggested (which is moving the head down and up in average 2 times).

Moving to the “No” answer from the results (Table 4), 100% of the participants used the same gesture to perform the negative answer, by moving their heads from side to side (also 2 to 3 times) and the average amount was also 2 reciprocations among more than 50% of the participants. Comparing these results to the assumptions table (Table 3), there was no participant using the first gesture. All 10 participants used the second gesture (which is moving their heads from side to side in average of 2 times) to answer “No” in reply to the questions. Therefore, 100% of the participants performed the second gesture.

Regarding the answers in reply to the questions that participants didn’t have an absolute answer to give and replied with something like “Sometimes”, “Maybe”, “Almost” and “I don’t know”. Participants performed 3 different gestures. 5 participants used the predicted gesture similar to the gesture form assumptions table (Table 3). They tilted their heads from one side to the other side from 2 to 3 times. One other participant used the same gesture as the previous 5 but with very slight movements. 3 other participants tilted their heads to one side and lifted their shoulders up. The one remaining participant used the same gesture that was used for answering “No”.

Considering these results for the “Maybe/Almost” answers, about 55% to 60% of the results could be considered as predicted (Table 3), meaning tilting the head from side to side in average 2 to 3 times. About 30% of the participants tilted their heads to one side and about 10% used a gesture intended for other types of answers.

5.2 Results for Finding the Sound’s Source experiment

Results from the second experiment consist of two series of directions guessed by participants in short form of F = Front, R = Right, L = Left, B = Back and U = Up.
Considering the data from the static sound experiment, for sounds positioned on “Right” side of participants’ head, all users guessed the correct location of the sound and the success rate is 100% for the “Right” position.

For sounds positioned on “Left” side of participants’ head, 8 out of 9 participants guessed the location as “Left” and one as “Up-Left”. This translates to a success rate of 94.4% for the “Left” position.

Regarding the sounds positioned in “Front” of the participant, only one out of 9 participants guessed the correct location. Another participant guessed “Front-Up”. The success rate was 16.7% for the “Front” position.

Moving to last position, 6 out of 18 “Back” positions (in general) guessed correctly by 9 participants and on “Back-Left” location was guessed as 50% correct location. Considering this, there was a 36.1% success rate for the “Back” position for static sounds.

The next category of spatial sounds are dynamic. For the “Right” positions, 14 out of 18 sounds positioned on right side were guessed correctly. The 4 other guesses were 50% correct including “Right” in combination with one other direction. That defines the success rate for the sounds positioned on “Right” hand side as 88.9% when the sounds were dynamic.

For the “Left” hand positioned sounds, 7 out of 9 guesses were correct and other 2 were 50% correct. “Left” in combination with another direction results in 88.9% success rate for dynamic sounds positioned on “Left” hand side of the participant’s head.

On the “Front” position, 6 out of 8 guesses were correct. One participant couldn’t say which direction the sound was positioned on. Considering that one blank answer as a wrong guess, 6 out of 9 guesses were correct, which results in a 66.7% success rate.

Finally, for the sounds positioned behind the participant (called “Back” in the data gathering process), 14 out of 18 positioned sounds were guessed correctly by participants, which results in a success rate of 77.8% for dynamic spatial sounds positioned on the “Back” side of the head.

Regarding the statistics, the average success rate for static sounds on all 4 positions is 61.8% and the average success rate for dynamic sounds on all 4 positions is 80.6%.

The average time for guessing the static sounds’ position has been 2.89 seconds and for dynamic sounds, it has been 4.31 seconds.

<table>
<thead>
<tr>
<th>Sounds Type</th>
<th>Direction</th>
<th>Front</th>
<th>Back</th>
<th>Right</th>
<th>Left</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Sounds</td>
<td></td>
<td>16.7% (1.5 out of 9)</td>
<td>36.1% (6.5 out of 18)</td>
<td>100% (9 out of 9)</td>
<td>94.4% (8.5 out of 9)</td>
</tr>
<tr>
<td>Dynamic Sounds</td>
<td></td>
<td>66.7% (6 out of 9)</td>
<td>77.8% (7 out of 9)</td>
<td>88.9% (8 out of 9)</td>
<td>88.9% (8 out of 9)</td>
</tr>
</tbody>
</table>

Table 5. Success rate for guessing the 3D positioned sounds.
5.3 Results for Trying a Sonic-only Game

The data collected for this experiment was based on the time that users spent before being able to dodge the laser beams shot towards them.

9 participants tried the game after a short explanation how to play the game. 2 participants had played the game previously. The two players who had already played the game could manage to dodge the laser from the first shot and had no waiting time before their first save. For the other 7 participants it took from 5 to 40 seconds to get used to the game and interacting with the system.

In total, participants had an average of 67.4% success rate in avoiding the laser beams and a 32.6% failure rate.

The minimum amount of saves was 3 out of 16 and maximum amount of saves was 15 out of 16 with minimum 0 seconds before the first save and maximum 40 seconds out of 160 seconds game play. Reviewing these statistics, the average seconds before the first save and getting used to the interaction was 8.9 seconds and the average amount of saves was 10.8 out of 16.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time before first save</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>Success rate</td>
<td>32.6%</td>
<td>67.4%</td>
</tr>
<tr>
<td>Number of saves</td>
<td>3</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 6. Statistics from Sonic only Game.

6. Discussion

Of course there is still a lot left for the future work and this is just the beginning. I believe that more research about combining different types of embodied interaction with existing interactive systems (and especially sonic interactive systems) will lead to better interactive systems in the future. As a result, I predict more user centered design and user friendly UX in the future.

If we look back at the experiments conducted, there are also several results not directly relating to the research questions previously posed. For example, if we look at the difference between static and dynamic sound environments, there is a clear discrepancy when it comes to the ability to determine sound direction. This issue of “flat sound” can be an important consideration when designing systems reliant on sound.

One of the results not discussed further, but which could be an interesting research question in future experiments, would be to determine a model for increased recognition of “Maybe” or “I don’t know” kind of answers. The results posted in this thesis indicate that there might be personal and/or cultural differences that prevent for proper recognition. Further research would be necessary to study how it would be possible to increase the odds of proper recognition of head gestures other than “Yes” and “No”. In a more permanent installation, it would be
possible for users to “calibrate” their answers to allow for more reliable answers – something to take into consideration for future experimentation.

There is also more room for investigating the feasibility of the sonic gaming, especially for visually impaired people.

Also while finalizing this paper, I also created another prototype for a hackathon completion held by “Music Tech Fest” group (http://musictechfest.org/) in Umeå, Sweden. By this prototype I examined the possibility of combining the user’s head movements and sonic interaction for playing music.

The prototype works in the way of tracking user’s head movements and detecting the taps which user produces by moving their head forward and backward. This taps will be timed and converted to a tempo as BPM (Beats per Minute). This BPM value then will be used to select and play a song with a similar BPM from user’s local or online music playlist.

The experience was very successful and the prototype was very well considered by musicians and experts who attended the competition. These musicians and experts who I discussed the ideas with, think that these types of interactive systems would probably be very effective and useful and they will be interested in using such interactive systems.

Also, my whole study is funded and I will have the opportunity to research and develop my ideas even further; Thanks to the #MusicBricks (http://musictechfest.org/musicbricks/).

7. Conclusions

In the Design section of this paper, we created three more detailed questions in relation to the main research question for this study. As a reminder, the main question is “How can Embodied Interaction be used to Improve and Enhance Sonic Interaction?”

Also, the three more detailed questions are:

1. Could embodied interaction in the form of head movements be used to interact with a Sonic Interactive System?
2. Could embodied interaction in the form of head movements be used to improve Spatial Sonic Interactive Systems?
3. Could embodied interaction in the form of head movements and Sonic Interaction be used for developing new methods of gamification?

Considering the results from the related studies, there are indicative results to answer all 3 research questions regarding the combination of the “Embodied Interaction” and “Sonic Interaction”, called “Embodied Sonic Interaction” in this paper.

The results from first experiment and also the afterwards interviews with users revealed that, by combining the “Embodied Interaction” and “Sonic Interaction”, there are possibilities to create new, more natural feeling interactive systems for users. In this case, there is a strong potential for using head gestures as way of interaction with sonic interactive system for confirmation or denial of the commands and questions coming from the interactive system just by tracking user’s head movements and recognizing “Head Gestures” when no verbal interaction with the system is possible. In other words, a wide range of users from all around the world confirms a command or question by moving their heads up and down and deny them
by moving their heads from side to side and interactive systems could easily benefit from this “Embodied Interaction”.

Results from the second experiment reveals that “Embodied Interaction” could enhance a “Sonic Interactive System” (which uses spatial 3D sounds to simulate a virtual environment to represent the correct position of the sounds). The performed studies indicate that, by combining the user's head movements with a spatial interactive system, the spatial hearing for the users could be enhanced by 18.8% in general, 50% for the sound sources positioned in front and 41.7% for the sound sources positioned behind the user in the spatial interactive system, making the system more reliable and usable.

Finally, the results from the last experiment show that a “Sonic Interactive System” (which has been designed in an embodied way and interacts with user's head movements), could form a real “Sonic Game” that is independent from any form of visual interactive interface. Such interactive system is also both feasible, in context of interaction, and enjoyable to play according to most users’ opinion. This type of interaction, and more specifically “Sonic Game Plays”, could also be suitable for people who are visually impaired, as it is sonic only and not relying on any other kind of visual interactive interface.

As a conclusion to all three research questions and in answer to the main research question, embodied interaction combined with sonic interactive systems, specially tangible and wearable interactive systems can enhance the user experience by delivering new ways of communication and interaction between users and interactive system.

This combination could also be used to improve sonic interactive system's user experience by reducing the failure of delivering the desired information to the user; For instance, it could be used for improving the artificial 3D sound experience.

Last but not least, embodied interaction in combination with sonic interaction could be used to create new types of interactive systems like gaming and similar interactive systems as an enhancement to sonic interaction user experience.
References


