Interaction with sound for participatory systems and data sonification

Liu, D.

CHAPTER 3

The Development of an Interactive Installation using Sonification: Bǎi/摆

This chapter is partially based on the following publication:

Liu, D., Kroese, J., & van der Heide, E. (2018). The development of bǎi: An oscillating sound installation. In *Interactivity, Game Creation, Design, Learning, and Innovation*, (pp. 69–79). Springer.

3.1 Introduction

In Chapter 2, we have concluded an ideal sound interaction framework as a responsive dialogue between the audience and the performance (cf. definition 2.10). While the audience participation and interaction starts the dialogue at lexical level, the semantic level of this dialogue involves auditory feedback and influences the audience's comprehension. A sound design complements this dialogue at the syntactic level by assisting the audience to learn and understand the system. We found such dialogue could be developed through a loop between the progress of learning and interaction (cf. Figure 2.4). This may achieve the possibility of a more continuous interactive system we proposed in section 2.5 for audience participation. We intend to apply the ideal framework to the interaction design of a sound installation, to examine whether an installation can constantly engage the audience through the interaction (cf. Figure 3.2).

As a starting point of such installation we would develop a musical instrument for audience participation. Hereby we first define a musical instrument as:

Musical instrument

Definition 3.1 *A musical instrument is a device which can be controlled to produce musical sounds.*

Since the invention of the loudspeaker, investigations, composers and artists have explored various ways of using speakers ranging from multi-channel speaker setups and hemispherical speaker designs to speaker sculptures and wearable speaker-based instruments. While speakers are often used in static positions, Gordon Monahan's *Speaker Swinging*, first performed in 1982, did apply a moving speaker as a *musical instrument* in a live performance. Three performers each swing a loudspeaker in circles with a sine or square wave as source signal (Monahan, 1982). The resulting sound is subject to the Doppler effect and the acoustic properties of the space. In 1968, Steve Reich pioneered the pendulum principle in his *Pendulum Music* (Reich, 1968). The performance involves phasing feedback tones resulting from suspended microphones swinging above the speakers (cf. QRcode 3.1). *Spatial Sounds (100dB at 100km/h)* by Marnix de Nijs and Edwin van der Heide (2000, 2001) is an interactive installation using a moving speaker. The installation interprets the position and movement of a visitor and reacts to it both with its movements and the real-time generated

QRcode 3.1

sound. In return, the visitors react to the installation and go through different experiences and emotions (van der Heide, 2010).

What these works have in common is that they exploit the physical properties of a moving sound source ,or microphone, in their design. The development and possibilities with loudspeakers are interesting and they provide a good basis for interaction and dialogue. However, more physical interaction can be added to the design. Therefore we will further investigate this in an interactive sound installation. We translate some of the ideas of Reich to this setup and investigate a moving speaker as the basis of a pendulum (cf.Figure 3.1). We decided to comply the moving sound source with our framework and use a speaker setup to achieve a responsive dialogue (cf. Figure 3.2). Additionally, we intend to investigate a motion data sonification design to determine whether the audience can understand the relation between the sound and the movement.

This chapter describes and reflects on both the technical and artistic decisions that were made during the design and development of an interactive sound installation, Bǎi. It covers the design goals and a short reflection upon what we have achieved so far.

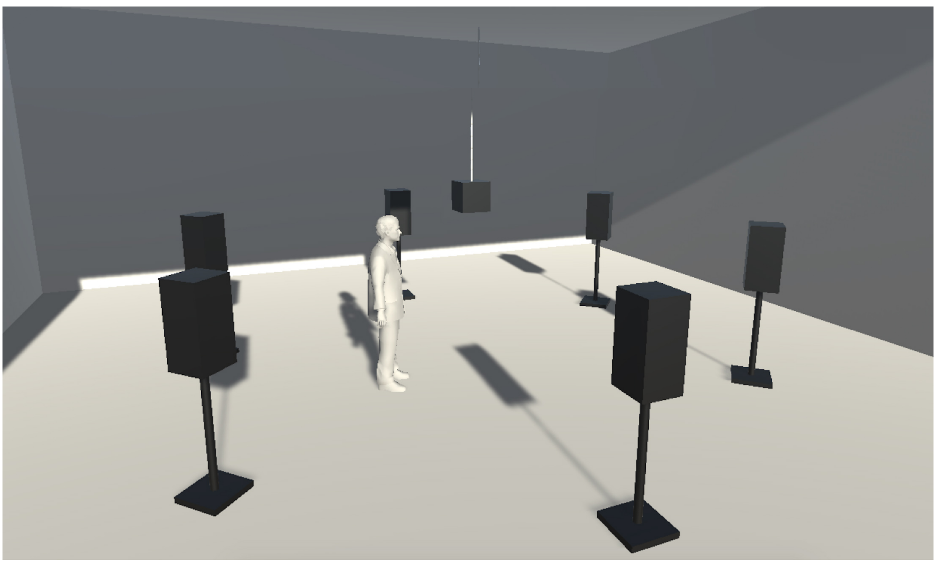

Figure 3.1: 3D model of the space setup for Bǎi/摆: an Interactive Sound Installation.

3.2 Interaction Design for Participation

In our approach we interpret the term interaction as a dialogue between the audience and the installation (cf. definition 2.1). We aim at developing the interaction and behaviour of the installation as surprising and intuitive.

Surprise

Definition 3.2 *In a surprising dialogue the two parties communicate and react to each other while neither of the two parties is fully predictable, nor has full control over the situation.*

Intuitive

Definition 3.3 *In an intuitive installation, the interactive form is easily understandable, so that the audience is able to explore the system without specific instructions.*

The interactive behaviour is not static but develops in order to realise an interesting ongoing dialogue. We have set a number of goals to help us achieve this:

1) use analogies between the physical input and sonic output of the system,
2) give the audience the experience of interacting with a system that reacts to their input but also surprises them with its own unpredictable behaviour,
3) make the audience aware that their actions impact the way the system behaves, without being able to fully control it,
4) make the audience perceive the system as 'beautiful', but also 'upset' or 'dangerous' through the changes of its behaviour (see section 3.3).

At the core of the interaction design is to experiment the ideal framework proposed in Chapter 2 in an interactive sound installation. We started with a moving speaker in the form of a pendulum as an interface for the audience to actively interact with (cf. definition 2.6). The pendulum speaker is suspended from the ceiling, surrounded a 6-speaker setup standing on the floor (see Figure 3.1).

Through pushing, pulling and rotating the pendulum speaker, the audience can set it into different oscillating motions (see Figure 3.2). We choose for the installation not to have a fixed form of interaction but rather have it alternative between different rules, and therefore different modes of behaviour (cf. in section 3.2.4). Furthermore, we have chosen a stochastic system for sound production in which the surrounding speakers react to the pendulum as it approaches them resulting in a dynamically changing sound environment. At first, it may

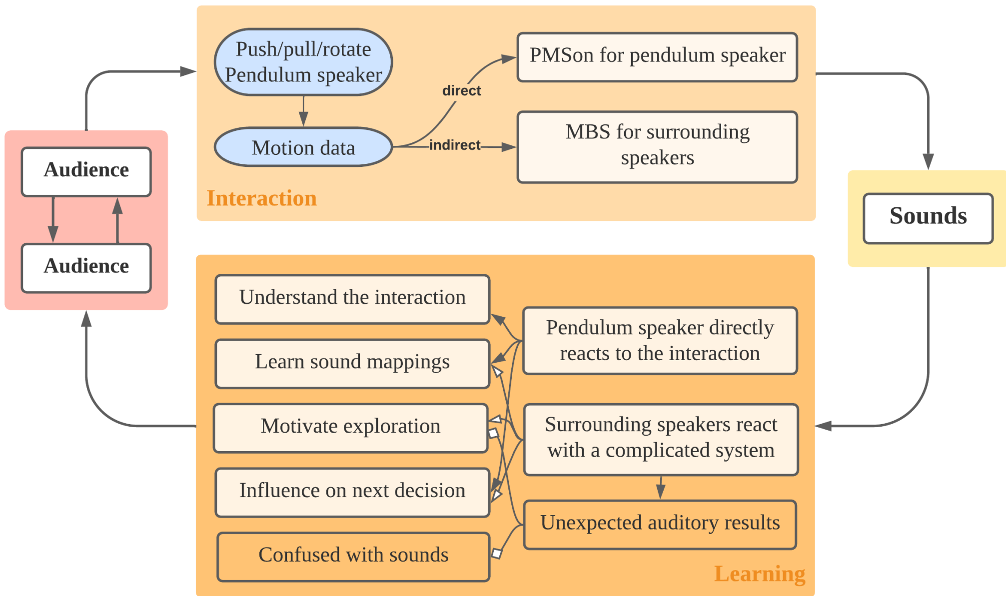

Figure 3.2: Ideal Framework for Băi. The color code refers to the elements of the dialogue model: subject-audience, verbal-actions, adjective-sounds.

seem that the environment reacts to the motions in a predictable manner. However, the self-movement of the pendulum influences the behaviour of the system, even when the audience does not directly interact with it; this brings unforeseen results. Combined with the fact that physical labour is needed to restrain the pendulum, this leads to a tense dialogue between the participant and object, struggling for control.

3.2.1 Technical Requirements

First of all, the position of the pendulum needs to be known to the system. An HTC Vive base station is mounted on the wall inside the room, and emits infrared signals. An HTC Vive tracker is placed on top of the pendulum speaker, in order to continuously collect the absolute position and orientation data of the speaker in the room. The data is transmitted to a computer running a patch in Pure Data (Pd), a real-time graphical programming environment for audio and graphical processing [1]. In Figure 3.3, we show how we have programmed

[1] Pure Data, `https://puredata.info`

the sensor interpretation, the rules for the interactive behaviour and the real-time sound synthesis for the surrounding speakers in Pure Data (version 0.50). Furthermore, Pure Data is controlling a software synthesizer in Ableton Live (version 9) for the sound generation of the pendulum speaker.

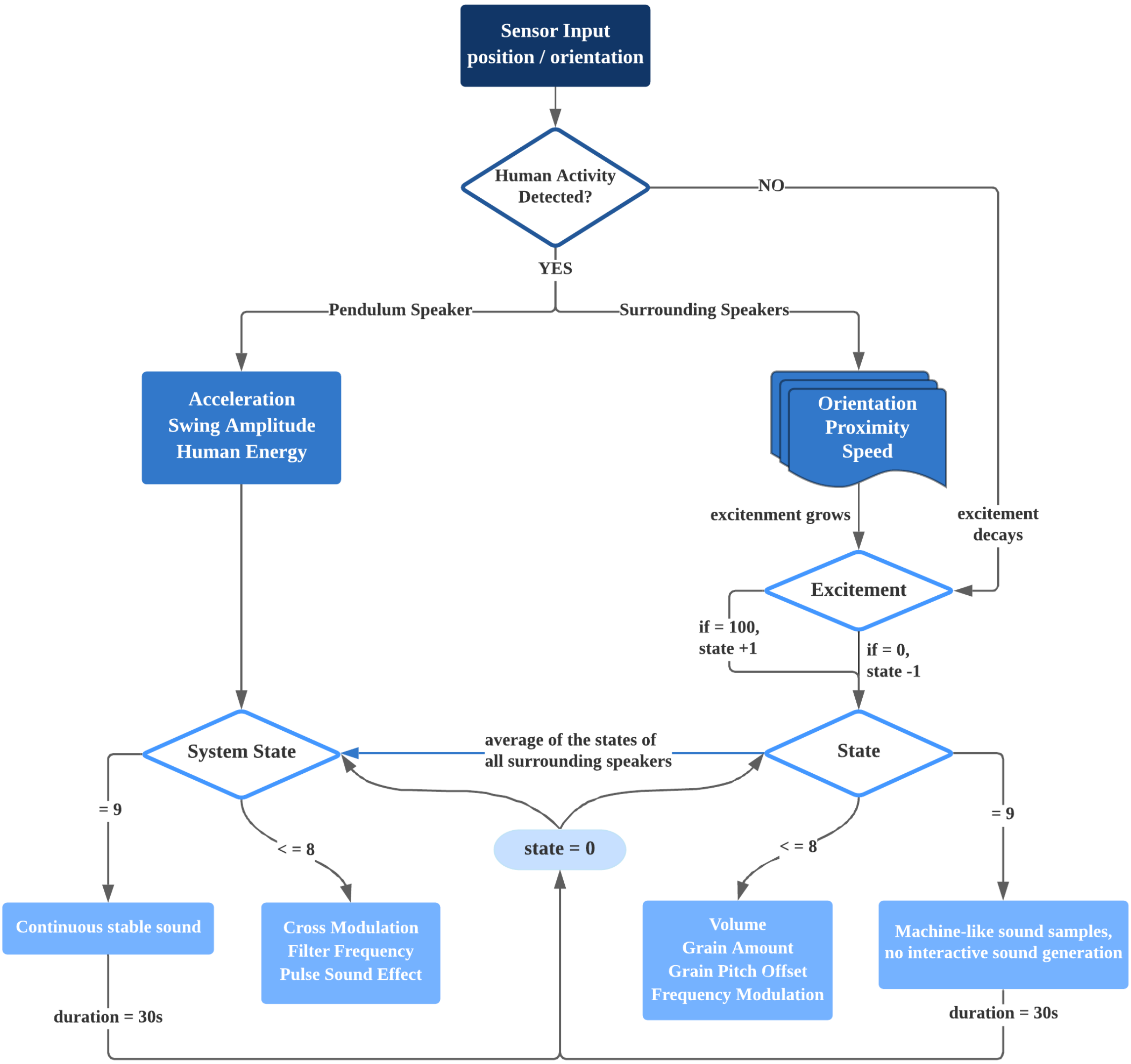

Figure 3.3: Process workflow of software development.

3.2.2 The Pendulum Speaker

From initial and preliminary experiment, we have observed that the pendulum itself has a strong and clear form of, what we would call, natural behaviour. It is a clear inherent physical behaviour, an oscillating motion slowly decaying because of the friction with the air. This makes interacting with the pendulum speaker, and thereby the installation, not so much a process of having full control over the system, but rather a process of using and directing the behaviour of the pendulum. The audience can interact with this behaviour by accelerating, holding and rotating the speaker. It can swing in linear or an ellipsoid orbit. After moving the speaker, it will continue to oscillate corresponding to the new energy applied to it. We decided to try to distinguish the natural motion of the pendulum from the audience interacting with it. In order to do this, we have developed an algorithm that learns the period, phase and amplitude of the swinging behaviour, we then analyze and compare the current phase and position of the pendulum with the predicted natural movement. This way human interruptions of the natural movement can be immediately detected and its energy can be quantified by calculating the amount of deviation. The detected human energy put in to the installation is used to influence the sonic and interactive behaviour of both the pendulum speaker and the surrounding speakers (see Figure 3.3). After interacting with the speaker, the algorithm stores the new swing movement and interprets it as the new natural movement. We believe that this direct form of interaction, realised in this way, gives the audience a feeling that the pendulum is alive and able to respond to the audience's actions. We intend this to result in a playful and physically intensive interactive endeavour.

In order for the pendulum speaker to produce swings that would not move too fast or tilt too high, the length, weight and mounting point of the cable that holds the pendulum are important design parameters. We established a minimum cable length of 3 meters. The mounting point of the cable is placed 0.5 meters above the pendulum's centre of weight, to keep the speaker relatively stable.

3.2.3 The Space

For the first presentation of the installation we chose to use six surrounding speakers placed in a circle around the pendulum (see Figure 3.1). Each of the surrounding speakers is functioning as a separate entity that individually reacts to the movement of the pendulum. It is designed in a way that a surrounding speaker reacts when the pendulum moves towards it. Since the pendulum can swing 360 degrees in the horizontal plane, the arc length or the distance between two neighboring speakers cannot be too wide so that the audience can clearly recognise the interaction between the pendulum and the surrounding speakers. Accordingly, the minimal amount of surrounding speakers is six since 60 degrees is an optimal angle for a small room setup. If the exhibited space becomes larger, it is recommended to add more surrounding speakers.

The installation is not meant to only interact with a single audience member. The swinging movement in space makes it possible for multiple audience members to interact with the installation in the same session. In that case, the audience does not only interact with the speaker but also interacts with each other through the installation. Furthermore, the audience can play different roles and alternate between engaging with the installation or just observing the progression (See Figure 3.5).

3.2.4 Software Development

One of our goals for the experience of the installation was to give the audience the feeling that they are interacting with a system that has a form of autonomous behaviour. The installation was designed to noticeably react to the audience, but also have a certain amount of unpredictability in how it will react. Furthermore, in order to motivate the audience to interact with the installation for longer periods of time, we chose to let the behaviour evolve as a result of the amount of energy that the audience puts into the installation.

Excitement and State

In our system, each of the surrounding speakers forms a separate entity that produces its own characteristic sound. The character of the sound is determined by calculating two main features for each speaker: 'excitement' and 'state'.

$$Speaker\# \, [E, S] = [0 - 100, 0 - 10]$$

These features were implemented in order to achieve an evolving form of interactive behaviour. The state determines both the character of the sounds that are produced and how the speaker reacts to the movement of the pendulum.

The level of excitement ranges from 0 to 100. It is a parameter to describe how 'excited' a speaker is within its current state. It is continuously updated by an algorithm that uses:

1) audience interaction — how much energy has been put on the pendulum speaker, which is measured by the acceleration differentiated from natural movement;

2) the proximity of the pendulum to the speaker;

3) the duration the pendulum is within a certain proximity of the speaker;

4) the speed of the pendulum;

5) the level of excitement of its neighbouring speakers.

The excitement level is calculated for each speaker separately. When a speaker's level of excitement reaches 100, it shifts to the next state. When its level of excitement decreases to 0, it falls back to the previous state. There are 10 states in total, starting from 0. Each state has its own mapping strategies for sound production (cf. section 3.3). They are designed in such way so that the audience perceives a clear change in sound and interactive behaviour and gets challenged by the new interactive behaviour.

We have added some additional rules to the state changes. A speaker can only switch to a new state when the state difference between the speaker and its neighbouring speakers is less than "3". Otherwise, it will not change state and influence the neighbouring speaker's level of excitement instead and wait for it to get to a higher state. Due to these rules, the system as a whole evolves as a result of the individual speakers' behaviour.

The pendulum speaker also has its state. It is determined by, what we refer to as the system state. The system state is the average of the states of all surrounding speakers. An important exception to these general rules is that when the system state equals to "8", all of the speakers' states will shift to "9 - chaos". This state lasts 30 seconds as a clear stage and builds up to a point where there is no return, because the feedback mechanisms in the system drive it into a state of uncontrollable chaos where all speakers stop being influenced by the audience (cf. QRcode 3.2). We will discuss this special state in detail in Section 3.3.

QRcode 3.2

Increase and Decrease of Excitement

The level of excitement increases while the pendulum is close to a surrounding speaker and it slowly decays while the pendulum is further away from it, this is modelled with parameters of growth rate and decay rate. The growth rate of the excitement varies with the amount of audience interaction. More interaction results in a higher growth rate. When the pendulum is following its natural movement the growth rate will start to decrease and the speakers start ‘cooling down’. Each surrounding speaker has its own growth rate related to the audience interaction. We expect that the implementation of growth rate would add a responsive nature to the system. We have decided to make the decay rate increase once a surrounding speaker reaches state 4, which means its level of excitement will decline faster and its state will easier fall back to its previous state. Thereby it becomes harder for a surrounding speaker to reach a state higher than 4, especially when there is no continuous human activity detected by the pendulum. We assume this rule would help to keep the audience interacting with the installation. To avoid that the states would alternate too fast, a minimum time that a state lasts has been defined.

3.3 Sonification Design for Motion Data

The participation experience is largely depending on the sound design, i.e. adjective element, as this comprises the syntactic level of the dialogue. We do not want the pendulum speaker to only act as an interface for triggering sounds in the surrounding speakers but intend to create a responsive performance model (cf. definition 2.15) for its interactive and expressive behaviour. In other studies, for example, Livingstone and Miranda introduced the term 'responsive sound environment' as *“a system that regenerates a soundscape dynamically by mapping ‘known’ gestures to influence diffusion and spatialization of sound objects created from evolving data”* (Livingstone & Miranda, 2004). In order to achieve the responsiveness, the surrounding speakers not only react to the movement of the pendulum speaker but the pendulum speaker also expresses its own movement in its sound and clearly reacts to people touching and moving the speaker.

To that end, we have chosen to use two different sound synthesis techniques to make a clear sonic distinction between the pendulum and the surrounding speakers. The pendulum generates machine-like (i.e., low to mid frequencies) sounds, whereas higher frequency sounds are generated from the surrounding speakers. PMSon (cf. definition 4.3) and MBS (cf. definition 4.4) are used to interpret and translate the input data into sound. These mappings make the installation react both directly and indirectly to the interaction with the audience. This gives the audience a sense of control, but at the same time makes the sounds resulting from the interaction somewhat unpredictable. Meanwhile, the sound results of the installation can develop from calm and peaceful to chaotic and aggressive. This was chosen to make the audience perceive the installation as beautiful and calming when handled with care, but also dangerous and distressing when handled aggressively.

3.3.1 Sound from the Pendulum Speaker (PMSon)

We use U-he Diva, a virtual analogue synthesizer in Ableton Live 9, to generate the sound for the pendulum speaker [2]. The control parameters are calculated in Pure Data and sent to Ableton Live via MIDI. The machine-like sound is produced by two oscillators passing through a voltage-controlled filter (VCF). Using cross-modulation, a sawtooth oscillator and a sine wave oscillator modulate each other's frequencies. The VCF is a low-pass filter, that filters the sounds of both oscillators.

As we have discussed in Chapter 2, *SoundBounce* (Dahl & Wang, 2010) successfully used a physical metaphor and PMSon design to realize an intuitive interaction with sound. Similarly, we use a metaphor of a mechanical machine and parameter mapping approach (PMSon) for the sound design of the pendulum speaker, in order to provide an alternative of direct contribution (cf. definition 2.8) that is intuitive to recognize and understand. The control parameters for the sound synthesis are derived from the pendulum's own physical behaviour. To that end, the amount of human activity is mapped to the amount of the frequency modulation and the filter frequency (cf. Figure 3.4). The audience can 'power on' this machine by putting energy into the pendulum. The pendulum's position is used to play a single midi note that gets triggered every time the pendulum travels a specific distance in space. Furthermore, the amplitude of the pendulum swing controls the velocity of each midi note. When the pendulum

[2] Diva Homepage, `https://u-he.com/products/diva`

swings, it generates pulsed sound effects. The linear acceleration of the pendulum is mapped to cross-modulation between the oscillators, to make it sound like a machine engine that is operating and accelerating. Meanwhile, the rotation rate or the angular speed of the speaker is mapped to the pitch of the synthesised sound. The faster the pendulum rotates, the higher the sound. In this way, the amount of human activity is used to create direct auditory feedback when the audience interacts with the pendulum speaker. The more energy the audience is trying to put into the pendulum, the more active and powerful the machine will be, and the more dynamic the sound will be. The parameters gradually decline again when no one touches it. The state of the pendulum speaker is used to make it sound more aggressive. When the pendulum speaker reaches state 9, it stops triggering midi notes but generates a continuous and stable sound. The cross-modulation and low-pass filter are removed and the pitch goes much lower. The machine turns out to be 'over-excited', and cannot be controlled or influenced by the audience any more.

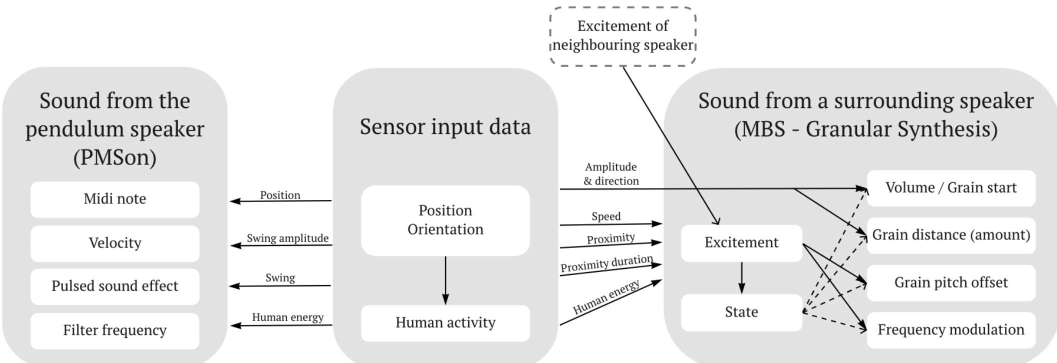

Figure 3.4: Mappings between the input data and the sound parameters.

3.3.2 Sounds from the Surrounding Speakers (MBS)

In contrast to the synthetically generated sounds from the pendulum speaker, the surrounding speakers produce a more natural sound. The sounds are generated using a granular synthesizer built in Pd. Each of the speakers have their own individual synthesizer that uses the same sound sample but with a pre-edited different pitch. The original sample is a recorded hit of a bell (cf. QRcode 3.3).

QRcode 3.3

When the pendulum hangs exactly at its equilibrium point in the centre of the space, the surrounding speakers will not generate any sound at state 0. But when the pendulum moves towards one of the surrounding speakers the sample of that speaker is played in full length. It sounds like the audience is using the pendulum speaker to hit the surrounding speakers, and 'awake' them.

As the interaction progresses, the granular synthesizer is used as a polyphonic sample playback engine in a later state.The distance between the pendulum and each of the surrounding speakers is mapped to the grain distance which sets the rate at which the grains are triggered and results in overlapping grains with a variable density. Currently, up to 100 overlapping grains can be generated resulting in dynamic and rich sonic textures. We have chosen this approach because we believe that this behaviour makes it intuitive for the audience to perceive what kind of effect the pendulum speaker has on each of the surrounding speakers. Next to that, the distance value is also mapped to the start point of each grain player. There is a clear and loud hit at the beginning of the sample. From observations we found that using the start point was an optimal parameter as opposed to the use of volume control since it applies the natural decay of the sound. The closer the pendulum moves towards a surrounding speaker, the louder sound it produces. This behaviour can be easily understood by the audience and it is intended to help them to understand the behaviour of the implemented excitement.

We have applied a model-based sonification (MBS) approach to implement different mapping strategies for the different states, in order to create distinct sonic characteristics for each state. Initially we implemented 5 states. As the state of a speaker increased, the sound transforms from stable harmonic tones into abstract and unrecognizable synthetic noise. However, this approach resulted in large and sudden changes between the states. Therefore, we decided to implement more states to be able to transform the original sample in a more gradual way. This makes it easier for the audience to perceive changes of the system while navigating through the various states. In order to create more complex dynamics in the playback of the grains, we have decided to add frequency modulation for states above 3 as well as to randomise the start point and pitch within a specific range of each grain. Consequently, the grains create a more complex, and use a wider sonic range. The original sampled sound gets dispersed because the hits are intensified and blurred as the speaker reaches higher states. The sound becomes more and more chaotic as the speaker gets excited. In state 9, all of the surrounding speakers play the full length of the original sample with frequency modulation and repeat at a random interval. The sounds become more machine-like compared to the sound in state 0, and assimilate into the pendulum's synthesised sound. After reaching the highest state the system 'cools down' and needs a little rest before it starts responding again starting in state 0.

3.4 Observations and Discussion

During a three-day exhibition at the NIME 2018 conference (see Figure 3.5), some observations of the audience interacting with the installation were made. We also had informal conversations about the work with some of the visitors. Although we did not use a strictly formal method for reviewing, our observations gave us some preliminary indications of how the audience reacts to, and interacts with the work. We use the same participation journey map (cf. Figure 2.1) to visualize audience experience at four stages (see Figure 3.6). We noticed that, at first, many visitors were mostly observing the installation instead of interacting with it. Some mentioned that 'they were not sure if they were allowed to touch the work'. After interacting with the installation, most of the visitors that we observed independently discovered the different forms of movement that the pendulum speaker reacts to, without the need for specific instructions. This seems to indicate that the basic form of interaction is intuitive and PMSon design may provide a more direct feedback for the audience to understand the interaction.

Most of the visitors also seemed to quickly notice that moving the pendulum towards a surrounding speaker resulted in this speaker reacting by playing a sound (see Figure 3.5b). Some visitors specifically mentioned that the interaction reminded them of handling a bell or wind chimes. Thus, it seems that the direct and noticeable sound results, as reference of direction contribution (cf. definition 2.8), can help the audience understand the interaction and navigate through different types of sound composition.

However, it seemed to not be clear in all to the audience that the sounds of the installation were able to develop from sounding calm to sounding aggressive and that the effects of the audience's interaction with the sounds would then also change. For some visitors, this was due to them handling the pendulum speaker so gently that the installation would always sound calm and not aggressive. Other

(a) Interaction 1: Gentle movement

(b) Interaction 2: Push hard

Figure 3.5: The audience interacting with Băi at NIME 2018 (Faces have been intentionally blurred to protect the privacy and anonymity of the individuals depicted).

(c) Experience for multi participants

(d) A contemporary dancer experimented with the speaker

Figure 3.5: The audience interacting with Băi at NIME 2018 (Faces have been intentionally blurred to protect the privacy and anonymity of the individuals depicted).

visitors did put enough energy in the installation to make it sound aggressive, but seemed to not be fully aware of how their actions altered the sounds. Although we also noticed that with the current setup, visitors needed some explanation before being able to experience the full dynamics and concept of the installation (see Figure 3.6). The design of the pendulum speaker might lack some obvious perceivable affordances. Moreover, a clearer distinction of sound design between different 'states' of the installation might help visitors to discover and understand the installation more easily. The MBS approach can achieve a more complex and indirect sound result (cf. definition 2.9). Lastly, our observations indicated that visitors had quite varying sensations while experiencing the installation. Some visitors avoided close proximity to the pendulum, but to others purposefully stood right under the pendulum to 'get a rush of it swinging right over their head' (see Figure 3.5c). Some visitors experienced the installation while laying on the floor and reported that it was a calming experience to them. This indicates that the audience was able to perceive both sensations of beauty and danger, which we aimed to convey with the installation.

A meta level of participation happened unexpectedly during the exhibition. There were two contemporary dancers who experimented choreography with the pendulum speaker (see Figure 3.5d). They gained understanding the rules of the

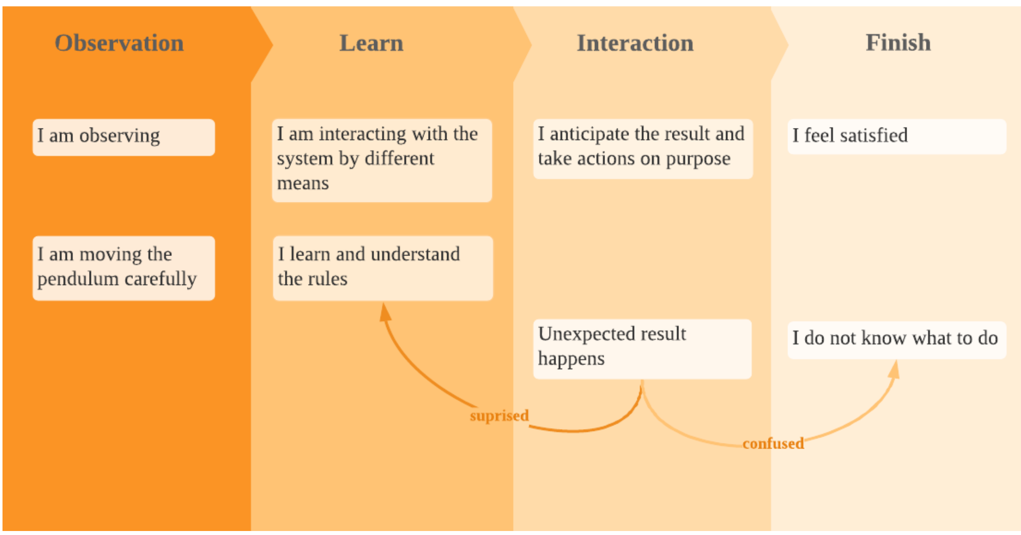

Figure 3.6: A participation journey map of Bǎi.

QRcode 3.4

changes of the sounds quick. Therefore, they moved the pendulum speaker in an intended artistic way and adapted their dance according to the changes in the sounds (cf. QRcode 3.4).

We believe that a good interactive installation should explain itself to the audience and establish a dialogue with the audience. In other words: it should steer the audience in such a way that it reveals its behaviour to the audience. Our initial observations indicate that there is still some room for improvement, but the original concept can be transferred.

3.5 Conclusion

We have developed Băi as a case study for the ideal framework proposed in Chapter 2. A pendulum speaker has been used as an expressive control interface by sensing its position, speed and rotation. Besides performing its own natural movement, the speaker gives both physical and auditory feedback to the sensed input. These aspects serve as a solid foundation for developing a dialogue between the audience and the installation. In this manner, we have achieved a dialogue model at semantic level, where the audience is able to pick up the clues that are provided by the system through sounds.

We have mapped out audience experience in a participation journey map to address how feedback assists the audience in learning and understanding the rules of the interaction (cf. Figure 3.6). We found PMSon approach is easier for the audience to recognise how the actions have a direct impact on the sound of the pendulum speaker (cf. Figure 3.2), which has contributed to the goal of achieving a good visibility of the auditory feedback.

Furthermore, we have developed the installation in a responsive way to challenge the audience when they are adapting their behaviour to the interface and interacting with sounds. While the audience are controlling and interacting with a system that noticeably reacts to them, the system also has its own behaviour and thereby a certain amount of unpredictability. Additionally, the pendulum speaker interacts with the surrounding speakers. While the movement of the pendulum speaker triggers sounds in the surrounding speakers it also influences the excitement and state of them. In return, the state of the pendulum gets affected by their state. We have constructed a dynamic relationship where the states of the surrounding speakers shift up and down, depending on the intensity and duration of the audience's input. On one hand, unexpected auditory results can surprise the audience and achieve continuous dialogue. On the other hand, they might be confusing and bring the interaction to an end (cf. Figure 3.6). In short, MBS approach establishes a stochastic sound system that is able to be developed constantly according to the interaction. While it can achieve a more dynamic and continuous interaction, unexpected auditory results may cause confusion for the audience and hence it takes more time and efforts to get understood (cf. Figure 3.2).

There are several options for the audience to engage in the installation (cf. Figure 3.5). One might just stand alone and observe the installation or walk around and move the pendulum speaker. It is possible for others to join simultaneously and either observe or join the interaction.

While our installation is based on a complex system, we have shown that the responsive interaction method (cf. definition 2.15) is easily understood, especially with the sound results that can be recognized and acquired. Meanwhile, the installation provides different levels of engagement depending on the interaction chosen by the audience. The initial observations of the prototype have given us some indications, but for a good review of the interaction of the audience with the installation, a thorough study would be a good addition to take this project to a next level.