Whisper: a web app for audience participation and multichannel diffusion

Treya Nash
Louisiana State University
treya.nash1@gmail.com

ABSTRACT

Whisper is a whispering gallery-inspired piece performed through a web app. The piece is designed for audience participation with cell phones. During the performance, audience members record themselves whispering into their cellphones, and upload their audio samples. The audio samples are continually loaded into
. A performer controls a separate webpage through which they diffuse the samples into multichannel space, using the JSAmbisonics library. The piece is intended to engage the audience and give them agency through the use of their voice and the choice of whether to upload or delete their recording. The soundworld of the piece is designed to create a highly intimate and immersive experience which plays with the intelligibility of speech. In the first section, audience members record themselves whispering lines of a poem. In the second section, audience members choose what they whisper, which is then obscured through granular processing using tone.js. The Whisper framework is designed to be adaptable for use in future performance pieces.

1 Introduction

Since 2001's *Dialtones (A Telesymphony)*, musicians have been incorporating cell phones and other portable devices into performance, on the basis of harnessing the ubiquity of modern devices to generate audience participation, engagement and novel ways of music making [4]. Key considerations in cell phone pieces include sound sources, audience interaction, and audience/performer relationships. Sang Won Lee discusses goals of ease of use, engagement, encouraging participation, and relationship between action and performance. [3] These ideas are central to *Whisper*. The app is designed to be easy to use, transparent in action and sound source, collective, and engaging.

The idea for *Whisper* comes from whispering galleries, spaces in which sound waves travel around curved surfaces, so that if a person whispers in one part of the space, the whisper can be heard clearly in another part of the space. One example is the whispering gallery in St. Paul's Cathedral. *Whisper* takes this phenomenon as a basis for ex-

ploration of intelligibility, anonymity, and collective performance. *Whisper* is designed to be performed in a multichannel space. The audience are prompted to record whispered words and phrases into their cellphones, which are diffused by a performer through at least four speakers using azimuth and elevation values.

2 Motivations

The motivations for this piece include audience engagement, exploration of space, collective music making, and the creation of a unique sound world. One of the most engaging features of using audience cell phones is the inherently spatial nature of distributing audio through their speakers. This distribution of sound causes cell phones to become a speaker array, as described by Benjamin Taylor [11]. *Whisper* incorporates the spatial distribution of cell phones to some extent, through the audience being able to play back their recordings before uploading them, but the primary spatial component of the piece is diffusion of audience whispers into a multichannel speaker setup. This is intended to create a kind of virtual whispering gallery, which has its own unique characteristics.

Within the framework of audience participation with cellphones, the choice of soundworld is vast. Many pieces transform the cell phone into an instrument, giving the audience the ability to play prespecified sounds as part of a collective. *Whisper* uses the voices of the audience members as the instrument, in an attempt to give the audience an authentic and individual contribution to the performance. The audiences' voices are sent to a server, and diffused by a performer into a multichannel space. The soundworld combines the whispering itself, the sound of the audience playing back their whispers, and the whispers diffused into the speakers. The intent is to create a highly intimate and delicate sonic experience, giving the audience agency in their recording process – they can choose whether to upload or delete their whisper – while preserving their sense of collective sound. While a whispering gallery is often used to highlight a single voice, *Whisper* is designed to obscure the individual voice, creating a unified sonic texture and protecting the anonymity of the audience. *Whisper* explores this idea through two separate sections. In the first section, the audience are prompted to whisper lines from Thomas Hardy's poem *In the Whispering Gallery*. These are diffused into the space without any processing. In the second section, the audience are prompted to record their own whispered thoughts. In the second section, their recordings are obscured using various granulation processes, so as to avoid

any anxiety surrounding the sharing of personal thoughts, and creating a new sonic layer.

3 Related Work

Whisper uses a similar framework to my 2022 piece, *Frog*, in which audience cell phones generate images of frogs with recorded croaks. In this work, as in Whisper, the audience connects to a server. A performer controls the sound sent to the house speakers, as well as the amount of frogs that can be produced. *Frog* is a highly collective experience, intended to turn the audience into a giant frog pond, while the performer controls the frog population. The aim of *Whisper* was to create a more personal, intimate experience, generating the sounds during the performance, and giving the audience a greater sense of agency.

Whisper draws inspiration from works that use cell-phone recorded audio sent and accessed through the web, such as Jesse Allison and Anthony Marasco's *Gravity Density*. This piece uses live audio from hacked CD players, which is sampled and edited by the audience, and sent back and forth between the audience and a server [**?**]. Whisper uses a similar system of recording and uploading audio files from cell phones to a server.

Auracle by C. Ramakrishnan, Jason Freeman, and Kristjan Varnik is a voice controlled collaborative web app which analyzes voices of remote participants and uses them to control a synthesizer [8]. Though *Auracle* and *Whisper* have very different interfaces and sonic outputs, they are both voice-driven experiences, designed to give the sound source a level of transparency and personal identification.

In the 2015 piece *Fields*, Tim Shaw, Sebastian Piquemal, and John Bowers create spatial relationships between audience cell phones and four speakers, using a combination of pre-recorded sounds and tone.js. This piece focuses on the distribution of sounds in space, using different granular synthesis parameters in different speakers as a form of spatialization [10]. Cell phones are turned into speakers, and instruments are controlled by performers. The structure of having audience generated sound controlled by a performer is used other cell-phone works including *Gravity Density* and *echobo* by Sang Won Lee and Jason Freeman. [3] The server design for both *Whisper* and *Frog* and draws from the NEXUS project at Louisiana State Univeristy [1], the Rhizome project[6], and the *soundworks* project at IR-CAM [9].

4 Technical Design

Whisper is written in JavaScript. The server is written using Node.js, and connections between the server and clients are created with socket.io. Tone.js is used for audio processing, and the JSAmbisonics library is used for multichannel diffusion. Since *Whisper* was inspired by whispering galleries, I wanted to create an experience in an immersive 3D space. The JSAmbisonics library allowed me to design a flexible system for multichannel performance [7]. Figure 1 shows the relationship between the audience, server, performer, and speaker system in *Whisper*.

The audience record mono files which are stored by the server. A separate webpage controlled by the performer accesses the sound files as they are uploaded, and sends them into the space. This can happen either through the random generation of azimuth and elevation values, or by using
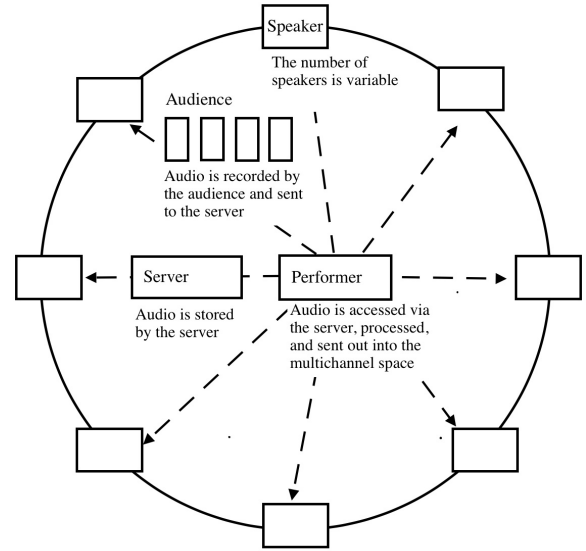

Figure 1: Technical setup for a performance of *Whisper*.

mouse movement to control panning. Figure 2 shows the page from which the performer controls the sound.

4.1 Audience Interaction

One of the key considerations in any music that involves audience participation is communicating the audiences' role without frustrating or alienating them. In the process of creating the concert piece *Poème Numérique*, Hödl, Bartmann, Kayali, Löw, and Purgathofer carried out an extensive study, "Large-scale audience participation in live music using smartphones". [2] As part of the study, they collected test data from potential audience members prior to the performance. In their findings, they discussed that while most audience members enjoyed the concept of the app, some cell phones failed to function as intended (detecting high-frequency audio signals), and some audience members felt their could have been more interaction. Other did not enjoy using cellphones in a concert setting.[2] This kind of study is highly useful in figuring out more effective practices for large-scale audience participation.

One of the main issues in designing pieces that involve cell-phone participation is audiences' expectation around detecting their individual contribution. In their paper on the *soundworks* framework and interaction topologies, Matusweski, Schnell, and Bevilacqua mention that the greater the audience, the less agency they may feel they have. [5] Works like *Poème Numérique* mitigate this issue somewhat by adding the visual component, in that case different colored lights on cellphones. *Whisper* endeavors to give the audience individual agency through their recording and quality control, but during the experience, they will not necessarily be able to pick out their audio when listening to the sounds through the speakers, due to the high volume of audio samples and the timbre of whispers. This is an area that needs further testing. In order for audience participation to function smoothly, careful design and testing is required. The user experience must be easy and intuitive. The success of a piece relies on the ability of the audience to almost in-

Figure 2: Performer interface in *Whisper*.

stantly understand the usage and function of an app. For this reason, the *Whisper* web app is designed to be as simple as possible. The audience are initially prompted to visit a url, which they can access through a qr code. This will lead them directly to the user webpage, through which they can record audio, play it back, upload it, or delete it.

In my initial testing of the *Whisper* app, audience members couldn't listen to their recorded audio. They were given a record button, and instructions to whisper into the microphone, but they were not able to hear their audio before uploading it to the server. I quickly realized that a lot of problems with volume and cut off could be avoided by allowing audience members control over their uploads. In the newest design, when the audience open the page they will see an information page giving them an idea of how the piece will work, seen in Figure 3. The performer can trigger the "Start Piece" button to appear whenever they are ready.

Once they click the start button, they navigate to the recording page, where they are given a list of instructions, seen in Figure 4.

When they press record, they are given a random line from Thomas's *In a Whispering Gallery* to whisper (Figure 5). The record button flashes red. I added the flashing to try and give a clear visual cue for when the app is actually recording. The text also changes to "stop recording" for added clarity. The feature of the poem line changing only as they start recording is intended to make sure they don't cut off the beginning of their whisper by accident. An envelope is added before playback.

After they stop recording, the play, upload, and delete buttons appear, seen in Figure 6. These use icons and colors to help with recognition. Allowing the audience to listen to their audio and choose whether to upload or delete it is an important part of the redesign which intends to give the audience further agency in music making. It allows them to do some quality control over their own audio, and delete it if they press "stop recording" too early, or need to adjust their distance from their phone microphone. Once they have either uploaded or deleted their audio, the piece resets,
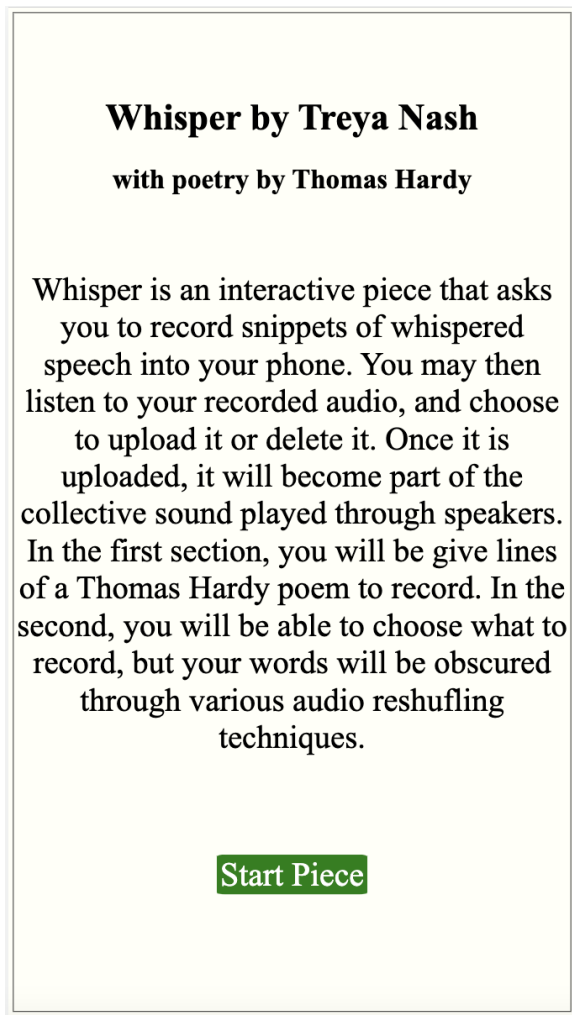

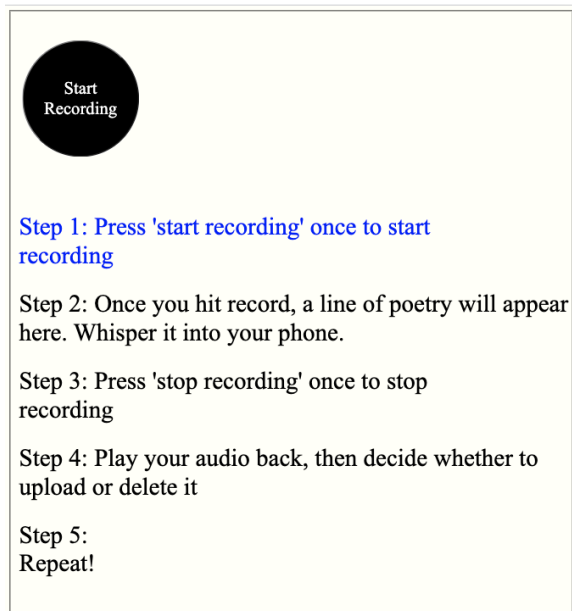Figure 3: Information Page for *Whisper*.
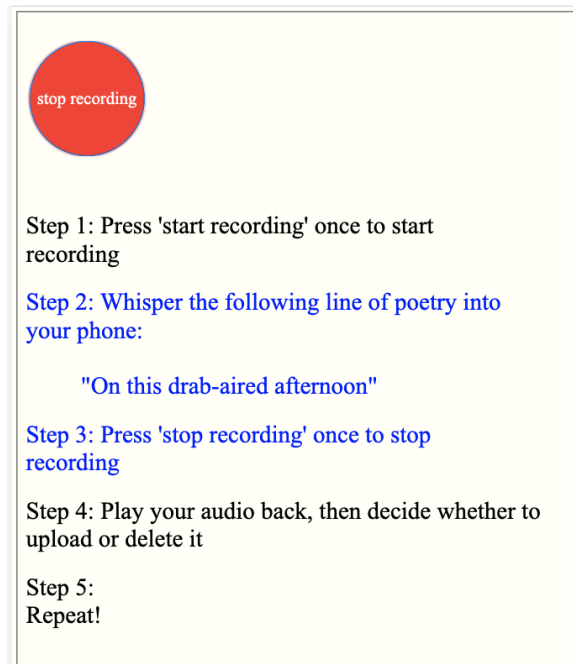

Figure 4: Instructions for *Whisper*.
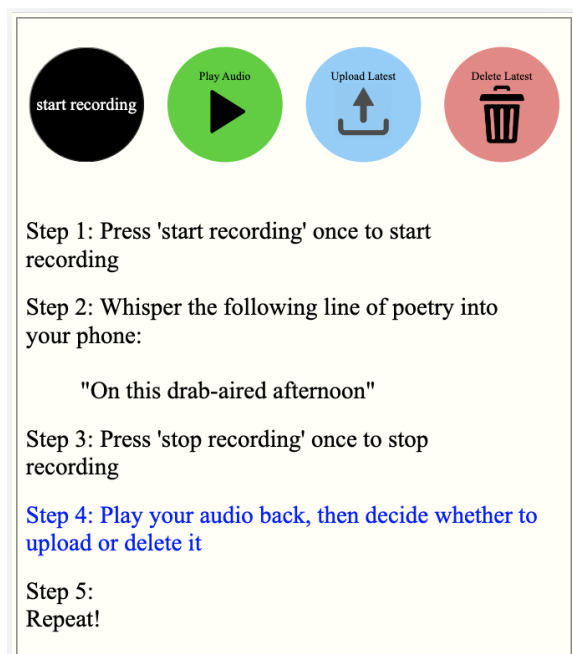
Figure 5: **Active Recording in** *Whisper.*

Figure 6: **After Recording** *Whisper.*

prompting them to record a new line of the poem and hiding the play, upload, and delete buttons. There is a cap to how many lines of audio each person can upload in the first section, and once they reach the cap the record button will disappear, encouraging time for listening as well as actively participating.

Allowing the audience to listen back on their own devices creates an extra spatial element to the piece. In its first iteration, the audience could only hear their audio coming from the speaker system. In this version, the audio is distributed between live voices, cellphones, and speakers, creating a multilayered spatial experience.

5 Conclusions and Future Work

Whisper succeeds in being collaborative, collective, and creating a clear relationship between the role of the audience and the sound of the piece. The user interface is still in the process of development, and will require further testing. My goal is to make it as intuitive and user friendly as possible, which requires as much input from general audience members as possible. The framework for recording and uploading sounds, and playing them through a multichannel setup, has a lot of potential, both for developing *Whisper*, and creating future pieces. The goal with *Whisper* is to create more complex spatial relationships, experimenting with creating predetermined panning paths, and perhaps using the general location of each audience member to inform spatialization of the sound. Some ideas that I have for future pieces include giving audience members an acoustic sound making device to record into their cellphones, and creating a piece for choir and cellphones, in which the choir record audio, and audience cell phones are used as speakers. In alignment with IoMusT concepts, one of the main goals of this project was to create a highly adaptable framework for audience participation and sound diffusion, which I will continue to develop and improve.

6 References

[1] J. Allison, Y. Oh, and B. Taylor. Nexus: Collaborative performance for the masses, handling instrument interface distribution through the web. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 1–6, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST.

[2] O. Hödl, C. Bartmann, F. Kayali, C. Löw, and P. Purgathofer. Large-scale audience participation in live music using smartphones. *J. New Music Res.*, 49(2):192–207, Feb. 2020.

[3] S. W. Lee and J. Freeman. echobo : Audience participation using the mobile music instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 450–455, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST.

[4] G. Levin, G. Shakar, S. Gibbons, Y. Sohrawardy, E. Semlak, G. Schmidl, J. Lehner, and J. Feinberg. Dialtones (a telesymphony), 2001.

[5] B. Matuszewski, N. Schnell, and F. Bevilacqua. Interaction topologies in mobile-based situated networked music systems, Mar 2019.

[6] S. Piquemal. Rhizome.
https://github.com/sebpiq/rhizome, 2014.

[7] A. Politis and D. Poirier-Quinot. JSAmbisonics: A
Web Audio library for interactive spatial sound
processing on the web. In *Interactive Audio Systems
Symposium*, York, United Kingdom, Sept. 2016.

[8] C. Ramakrishnan, J. Freeman, and K. Varnik. The
architecture of auracle: a real-time, distributed,
collaborative instrument. In *Proceedings of the
International Conference on New Interfaces for
Musical Expression*, pages 100–103, Hamamatsu,
Japan, 2004.

[9] N. Schnell, J.-P. Lambert, S. Goldszmidt, and
B. Matuszewski. Soundworks. https:
//github.com/collective-soundworks/soundworks,
2015.

[10] T. Shaw, S. Piquemal, and J. Bowers. Fields: An
exploration into the use of mobile devices as a medium
for sound diffusion. In *Proceedings of the international
conference on New Interfaces for Musical Expression*,
pages 281–284. The School of Music and the Center
for Computation and Technology (CCT), Louisiana
State University, 2015.

[11] B. Taylor. A history of the audience as a speaker
array. In *Proceedings of the International Conference
on New Interfaces for Musical Expression*, pages
481–486, Copenhagen, Denmark, 2017. Aalborg
University Copenhagen.