Displaced Listening

TAITO FUSHIMI, Faculty of Policy Management, Keio University, Japan

Additional Key Words and Phrases: Sound Walk, Sound Art, Urban noise, Sound-based navigation, Soundscape

1 Program Notes

Displaced Listening is a project set in Shibuya in which two strangers, each equipped with binaural microphones and earphones, exchange the sounds of their respective surroundings in real time. Relying on these exchanged auditory cues, they move through the city in an attempt to encounter one another.

As a center of consumer society, the city attracts large numbers of people and is constantly saturated with a wide array of sounds, including advertising audio from digital signage, street speeches, and public transportation announcements. In everyday life, however, many of these sounds are perceived as meaningless noise and are habitually ignored.

In this work, such sounds function as clues that bring the two participants closer together. Depending on each other's auditory perception, they navigate the urban environment and begin to discern the position and presence of the other within what would otherwise be considered noise. Here, noise is not treated as something to be eliminated, but rather as a medium that generates relationships with others. Through this shift in perspective, the project attempts to reexamine and reconfigure our understanding of the urban environment.

Fig. 1. "Displaced Listening", 2025.

Author's Contact Information: Taito Fushimi, Faculty of Policy Management, Keio University, Tokyo, Japan.

NIME '26, June 23–26, 2026, London, UK

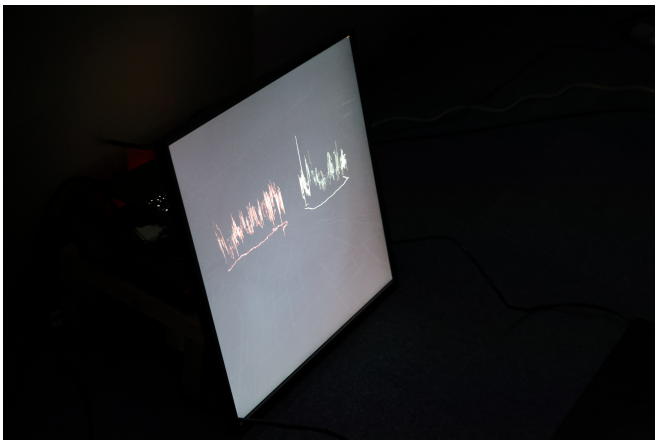

Fig. 2. "Displaced Listening", 2025.

2 Project Description

This project focuses on how urban environmental sounds that are heard but not actively listened to in everyday life can be perceived again. In urban environments, the most continuously present sounds are often mechanical in nature, including those of cars, advertising audio, trains, and jet aircraft. Although these sounds contain minor internal variations, they are generally experienced as monotonous and lacking in change[2]. Due to their pervasive presence, they recede into the background and are perceived as meaningless noise in daily life. This condition indicates that such sounds do not function as signals for individuals and are not recognized as information.

In this work, these environmental sounds are employed as cues that enable participants to encounter one another. Through this use, noise is transformed into information that actively guides behavior, in other words into signals. This transformation allows urban sounds, which are often flattened and perceived as lacking site specificity, to reemerge as perceptual objects grounded in a particular environment.

The work is realized through real time and bidirectional audio transmission using WebRTC. Participants wear binaural microphones and earphones and move through urban space while sharing their surrounding sound environments with each other. In this process, participants act not based on the sounds in their own immediate surroundings but on sounds from the location of the other. Through these sounds, they estimate the presence and position of one another.

Through this structure, the work expands the soundwalk from an individual practice of environmental perception into a relational form of navigation mediated by the perception of others. In addition, the sharing of sound reveals the presence of others who inhabit the same sonic environment, a presence that usually remains unnoticed in urban settings. In this sense, the work transforms sound from a mere object of perception into an interface that mediates relationships between people and enables the formation of new forms of connection within the urban environment.

2.1 System

This work is designed as a system that integrates real-time audio communication and location tracking in order to enable mutual exploration between participants mediated by sound within an urban environment. Two participants are connected through a custom-developed web application and move through the city while exchanging their respective surrounding sound environments in real time.

At the core of the system is a WebRTC-based audio communication mechanism that operates within a web browser. This enables low-latency, bidirectional audio streaming, allowing participants to listen to the other's "current sound environment" in a near-synchronous manner. Such real-time performance is essential for enabling orientation and movement based on auditory cues.

Each participant wears binaural microphones and earphones connected to a smart device. Binaural recording preserves not only sound intensity but also spatial cues such as directionality and distance, enabling the listener to more physically apprehend the sound environment at the other participant's location. This aspect plays a crucial role in the present work, in which sound functions as a means of navigation.

During each session, GPS data and audio recordings are simultaneously captured via the smart devices. The GPS data are used to visualize participants' movement trajectories retrospectively and to analyze the relationship between sound and behavior. The audio recordings, in turn, are stored for the purpose of revisiting the experience and for documentation, capturing aspects that may not be fully retained through real-time transmission alone.

Functionally, the system consists of four roles. The user1 and user2 interfaces correspond to the actual participants and provide minimal functionality focused on confirming the presence of the other and enabling audio exchange. The main interface is intended for experiment management and documentation, offering functions such as monitoring audio streams, tracking GPS trajectories, and storing data. In addition, the guest interface provides real-time monitoring functions similar to those of the main interface but is designed for observers and does not include recording capabilities.

In this way, the system technically supports the experience of exploring others through sound by integrating three key elements: low-latency audio communication via WebRTC, the reproduction of spatial listening through binaural recording, and the recording and analysis of movement via GPS.

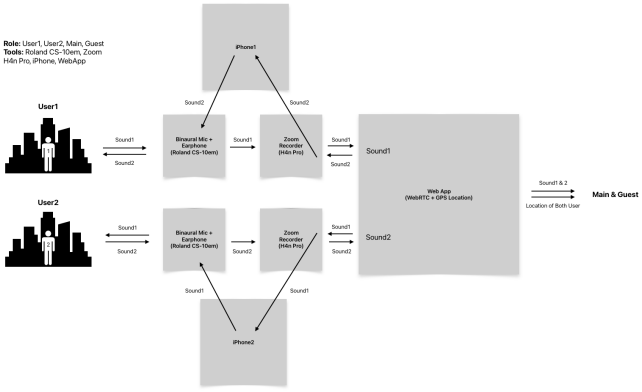

Fig. 3. System Overview

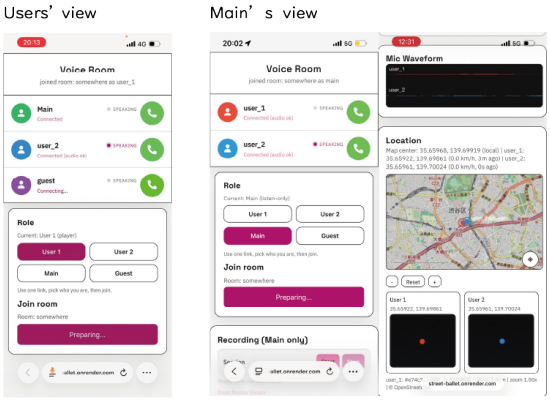

Fig. 4. Overview of The Web App

2.2 Observations and Reflections

This project was conducted through multiple sessions involving more than ten participants. Although encountering one another was not explicitly defined as a goal in each session, all pairs ultimately succeeded in reaching each other. This outcome suggests that sounds present in urban environments can function as effective cues for estimating the location of others and guiding movement toward them.

Participants' subjective reports further indicate that sound contributed not only to directional perception but also to the perception of distance. One participant noted that when brief and high intensity sounds such as ambulances or police sirens were heard through the earphones but not perceived in their immediate surroundings, this discrepancy heightened their awareness of the distance between themselves and their partner. In addition, when participants came into close proximity, urban noise was perceived as feedback or latency, which reinforced their sense of the other's presence.

These observations reveal that within this work sound functions not only as a cue for estimating the location of others but also as an indicator for perceiving interpersonal distance and proximity. In other words, environmental sounds that are usually not consciously attended to are transformed into meaningful information through relational interaction.

2.2.1 Navigation through Sound. One of the reasons all participants in this project were ultimately able to encounter one another can be attributed to the landmark-like properties of urban sound. In particular, persistent and high-intensity sound sources (e.g., street loudspeakers and station announcements) functioned as spatial reference points, analogous to visual landmarks.

Furthermore, the urban environment contains a multiplicity of sound sources occurring simultaneously, each with distinct characteristics. This enabled participants to estimate location by attending to differences among sounds. Participants adjusted their movement while tracking changes in the direction and intensity of these sounds, thereby using them as dynamic navigational information.

In this way, urban sound operates not merely as background noise but as identifiable and persistent spatial cues, enabling movement independent of visual reliance.

2.2.2 Site-Specificity: Shibuya. The realization of this work relies critically on the specific characteristics of Shibuya. Shibuya is not only a high density sonic environment but is also characterized by a fluid soundscape in which commercial audio, traffic noise, and human voices continuously overlap and shift over short time spans. In particular, in areas such as the Shibuya Scramble Crossing, human movement is closely coupled with the generation of sound, resulting in a dynamic acoustic environment that constantly changes in response to pedestrian flows.

Within this environment, sounds such as street speeches, advertising audio, station announcements, crowd murmurs, and intermittent traffic noise overlap while maintaining distinct temporal and spatial patterns. Participants are able to use these differences as cues to distinguish between the sounds in their immediate surroundings and those at the location of others.

Furthermore, the sonic environment of Shibuya is not a homogeneous mass of noise but is composed of a collection of locally differentiated sonic regions. This structure of sonic regions provides cues for spatial recognition and enables navigation based on sound.

Therefore, navigation in this work is not solely determined by system design but depends on the fluid and spatially segmented sonic environment specific to Shibuya. In this sense, the work is not universally applicable to any urban context but is grounded in an environment such as Shibuya, where pedestrian movement and commercial activity continuously reconfigure the urban soundscape.

Fig. 5. A image of Shibuya

2.3 Related works

The practice of the soundwalk has developed on the basis of R. Murray Schafer's concept of the soundscape and the field of acoustic ecology, as an approach that deepens spatial awareness and understanding of the social environment through the attentive listening of environmental sound[1]. In particular, Hildegard Westerkamp defined the soundwalk as "a walk

with the intention of listening to the environment," positioning it as an act of directing attention toward sounds that are ordinarily overlooked in everyday life[4][5]. Furthermore, in practical works, Janet Cardiff's audio walks transform the listener's perception of space by layering recorded sound onto real environments, while Christina Kubisch's Electrical Walks reveal otherwise imperceptible sonic environments by making audible the electromagnetic sounds embedded in the city[3][6]. These practices have played an important role in the rediscovery of environmental sound and the expansion of perception; however, many of them are grounded either in individual listening experiences or in precomposed sonic scenarios.

In parallel, within the domains of communication technology and networked systems, various attempts have been made to connect participants across distance and generate shared experiences. For example, Kazuhiko Hachiya's Inter Dis-Communication Machine reconfigures the relationship between self and other by enabling participants to exchange sensory information such as vision and hearing. In this work, the subject of perception is effectively displaced, producing experiences mediated through another's senses. Similarly, in the field of network music, represented by practitioners such as Chris Chafe, communication technologies have been used to share sound in real time across distant locations, exploring synchronous sonic experiences within distributed environments. For instance, the SoundWIRE project at Stanford University, as well as remote performances utilizing JackTrip, a low-latency audio transmission system, have enabled geographically separated participants to engage in simultaneous musical interaction via networks.

This work is situated at the intersection of soundwalk practices, which focus on the perception of environmental sound, and communication technologies that enable remote interconnection. That is, rather than treating environmental sound merely as an object of perception, or communication simply as a means of interaction, this project is characterized by its construction of a process in which participants move and approach one another through environmental sound itself as the mediating medium.

2.4 Relavance to NIME

This work is conducted within the high density and multilayered acoustic environment of Shibuya and is constituted as a practice grounded in this specific urban context. Sounds characteristic of the city such as street speeches, public transportation announcements, and commercial audio are typically perceived as background noise. In this work, however, they re-emerge as elements that mediate relationships between participants and structure a shared sonic experience. In this context, sound is not merely an environmental component or an object of perception but functions as an interface that mediates co-presence with others and generates relationships rooted in a particular environment. By employing real time networked audio and relational listening, the work engages directly with core concerns of the NIME community, including embodied interaction, distributed performance, and the design of new musical interfaces. It further suggests new possibilities for communication and community formation through sound, offering a novel perspective to the fields of soundwalk practice, sonic art, and interactive music systems.

3 Technical Notes

This work will be presented as a remote performance using a pre-recorded video approximately 15 minutes in length. It can be performed using standard stereo speakers and a screen or display. The video is an MP4 file with stereo audio and can be played locally.

The work does not require specialized spatial audio systems. However, clear stereo playback in a relatively quiet environment is recommended in order to preserve the perceptual differences in sound that are central to the piece.

4 Ethical Standards

This research and work comply with ethical standards. This research does not pursue financial gain or profit. Furthermore, prior consent has been obtained from all individuals participating in this project. This work aims to contemplate the modern city through sound.

5 Media Links

In this section, please provide a list of links to media representations of your submission.

- Trailer Video: https://youtu.be/fHcFTT6pd5o
- Performance Video: https://youtu.be/Ziv2EF_Oynw

6 Ethical Standards

All participants were fully informed about the project and provided their consent prior to participation. As the project was conducted in a public space, careful attention was paid to ensure that it did not cause disturbance or inconvenience to others.

References

[1] R. Murray Schafer. 1977. The tuning of the world. Knopf, New York.

[2] R. Murray Schafer. 1969. The new soundscape. Universal Edition, Vienna.

[3] Alan Licht. 2007. Sound art: Beyond music, between categories. Rizzoli, New York.

[4] Luis Velasco-Pufleau. 2023. Listening is action: A soundwalk with Hildegard Westerkamp. Performance Philosophy 8, 1 (2023), 86–100.

[5] Hildegard Westerkamp. 2017. The practice of listening in unsettled times. In Proceedings of Invisible Places: Sound Urbanism and Sense of Place, 29–45.

[6] Jennie Gottschalk. 2016. Experimental music since 1970. Bloomsbury Academic, New York.