Fragmented Voices, Activated Space: An Interactive Multimedia Installation Based on "A-Ronne"

MATTEO TOMASETTI, Conservatory of Music of Pescara, Italy
GIUSEPPE TAFFI, Conservatory of Music of Pescara, Italy
ENRICO FALCONE, Conservatory of Music of Pescara, Italy
MICHELE D'ALESSIO, Conservatory of Music of Pescara, Italy

Additional Key Words and Phrases: Interactive sound art and sound installations

1 Program Notes

Fragmented Voices, Activated Space is an interactive sound and light installation inspired by "A-Ronne" (1974) by Luciano Berio and the text by Edoardo Sanguineti on which the work is based. The project reinterprets Berio's exploration of voice, language fragmentation, and spatial dramaturgy through a real-time interactive system that links human vocal presence, sound diffusion, and light. A corpus of approximately 150 vocal samples was created from micro-fragments of Sanguineti's text. These short units, ranging from phonetic gestures to brief textual phrases, form a modular sound bank where voice is treated as acoustic-textual material rather than as a carrier of linear meaning. The installation was originally conceived for the foyer of the Music Conservatory of Pescara (Italy), where ceiling-suspended microphones and strategically placed LED bars create a hybrid scenographic and interactive environment. When participants speak into the microphones, the system detects vocal activity and triggers the playback of a spatialized vocal fragment in a quadraphonic setup, while simultaneously activating lighting events via Digital MultipleX (DMX) control. Each vocal gesture thus generates a coupled sonic and luminous response. The work explores the voice as an activating force that reconfigures space, producing an emergent acousmatic texture shaped by audience interaction and architectural context.

2 Project Description

This project is an interactive sound and light installation inspired by "A-Ronne" [2] (1974) by Luciano Berio, a collaboration between the composer and Edoardo Sanguineti, whose poetic text[1] forms the basis of the original work. Rather than quoting the composition itself, the installation reinterprets the relationship between voice, text, fragmentation, and spatial dramaturgy through an interactive multimedia system.

We recorded a large set of vocal samples derived from fragments of Sanguineti's text used in "A-Ronne". The recording sessions took place at the Music Conservatory and involved several singers and musicians who read selected phrases from the text. After completing the recordings, we edited the material, isolating micro-phrases and phonetic-textual units such as"nel mio principio", "nel mezzo del cammino", "della mia carne", "in my beginning", "is my end", along with other short textual fragments. The samples were then further processed through detailed editing, mixing, and manipulation in order to shape their timbral and expressive qualities. This process ultimately resulted in a sound bank of approximately 150 vocal samples. The voice is therefore treated not as a vehicle of semantic continuity, but as modular acoustic-textual material, echoing Berio's exploration of using voice as sound material.

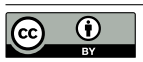
The installation was originally conceived for the foyer of the Music Conservatory building in Pescara, making the architectural and social space an integral component of the work. Several LED light bars will be distributed throughout the space, and four microphones will be positioned as visible scenographic elements.

In contrast to the initial design, where a single central microphone acted as the primary point of interaction, the system will be reconfigured so that all microphones will be simultaneously active. This change will encourage a more distributed and communal form of engagement, enabling multiple participants to interact with the installation at the same time.

Rather than guiding users toward a single focal point, the installation will support a multiplicity of interaction nodes, fostering a collective performative environment in which multiple voices will coexist and influence the system's behaviour.

[1] https://www.paulj.myzen.co.uk/blog/teaching/voices/files/2011/10/A-Ronne.pdf

Authors' Contact Information: Matteo Tomasetti, Conservatory of Music of Pescara, Pescara, Italy; Giuseppe Taffi, Conservatory of Music of Pescara, Pescara, Italy; Enrico Falcone, Conservatory of Music of Pescara, Pescara, Italy; Michele D'Alessio, Conservatory of Music of Pescara, Pescara, Italy.



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

The visual layout of the microphones, still carefully arranged within the space, will reinforce this shift from a singular to a shared interaction paradigm.

When a person speaks into the microphones, the system detects the presence of sound energy and, once a defined level is reached, generates a trigger event. This trigger simultaneously activates two parallel processes.

(1) it plays back one of the vocal fragments from the sound bank, spatialized over a quadraphonic loudspeaker system;
(2) It sends control data to a DMX lighting system, which activates specific light sources and lighting effects in the space (detailed in Section 4).

In this way, each vocal gesture from a participant produces a sonic and luminous response. The space becomes a responsive environment in which voice, sound diffusion, and light are tightly linked. The audience does not control specific sounds or meanings, but rather injects energy into a system that recomposes the vocal-textual material in unpredictable ways. The result is an emergent, ever-changing acousmatic texture shaped by the presence and activity of visitors.

Furthermore, the live voice of each participant was also routed through an additional processing chain developed in Ableton Live 12[2]. This processing stage introduced real-time transformations such as spectral filtering, time-based effects (delay and reverb), and granular-like diffusion processes that we detail in Section 4. Through these treatments, the participant's immediate vocal presence merged with the pre-recorded vocal fragments, extending the voice into an evolving sonic texture and reinforcing the acousmatic character of the installation.

Moreover, alongside the interactive system, we composed a 10-minute musical texture that functions as a continuous background layer. This material provides a structural and atmospheric foundation for the installation, while the triggered vocal fragments and live-processed voices occur as independent, superimposed events in real-time rather than being embedded within the fixed piece.

In addition, to facilitate the audience's understanding of the conceptual framework without requiring prior engagement with the installation, subtle visual cues will be introduced. For instance, excerpts from Sanguineti's original poetry will be projected onto the walls, providing a contextual layer that connects the audience to the work's textual and conceptual origins.

Conceptually, the installation extends Berio's investigation of the voice as a site where language, sound, and theatricality intersect [3]. Here, however, the compositional process is partially displaced onto the interaction between participants and the system. The work explores how fragmented text, disembodied voice, spatialized sound, and reactive light interact to construct a performative environment in which the audience's vocal presence acts as the catalyst for sonic, visual, and spatial transformation.

3 Positioning within NIME Research

The project situates itself within a strand of NIME research that moves beyond discrete object-based controllers toward spatially distributed and environmental interaction systems, where sound diffusion, audience presence, and spatial configuration become integral parts of the interface. The quadraphonic diffusion and synchronized DMX lighting system create an interactive audiovisual field in which space itself functions as part of the installation.

Furthermore, this installation contributes to the NIME discourse by shifting focus from conventional, object-centered musical controllers toward architectural-scale interactive systems that co-design sound, visual elements, and spatial experience. Recent NIME Installations proceedings document a growing interest in immersive and interactive spatial works, where audience presence, sonic distribution, and environmental components are integrated into the interaction design. Studies such as Birchfield et al.'s exploration of interactive public sound art emphasize the role of spatial and environmental considerations in interactive musical expression [1].

Moreover, interactive sound and light installations presented at recent NIME conferences (e.g., "near and far") illustrate how multimodal feedback can be driven by audience interaction within configured spaces [6].

This installation aligns with ongoing research trends in the NIME community in several ways. First, voice-driven interaction has been identified as a key area of interest in NIME research, with recent work mapping and categorizing voice-centered NIMEs to understand the variety of expressive and interactive paradigms explored in the field [3].

Second, interactive sound installations that respond to audience presence and real-time input have been regularly featured in the NIME Installations track, indicating community interest in systems that blur boundaries between interaction, sound, and space.

Moreover, recent projects presented at NIME incorporate not only sound but also visual or light feedback as part of the interactive system, demonstrating the relevance of multimodal environments for expressive control and audience engagement [4].

[2]https://www.ableton.com/en/live/

4 Technical Notes

The system is built around a real-time audio analysis and control architecture developed in Max 9[3] and Ableton Live 12. A Behringer UMC1820 audio interface handles signal routing. Participant vocal inputs will be captured through multiple Shure Beta 58A microphones, all of which will be active within the installation. The microphones, while serving as visible scenographic elements, will simultaneously function as distributed points of interaction.

The incoming audio signal from each microphone is processed by a threshold-based gate implemented in a custom Max for Live device (see Fig. 1). When the signal amplitude exceeds a predefined threshold, the gate opens and generates a discrete trigger event.

This trigger is routed to two parallel processes:

(1) The audio process controls sound generation through a custom Max for Live module hosted in Ableton Live, embedding a playlist of approximately 150 vocal samples described in Section 2. Each trigger initiates sample playback after a fixed latency of 2000 ms. This delay was intentionally chosen to allow participants to complete their vocal gesture and perceive the immediate processing of their own voice before the system responds with a triggered fragment. Sample selection follows a non-repetitive random algorithm that guarantees exhaustive traversal of the sound bank before any repetition occurs. At the end of each sample, an end-of-playback trigger is generated. The resulting audio signals are routed to a quadraphonic loudspeaker configuration, where spatialization is controlled through an algorithmic process implemented with low-frequency oscillators (LFOs) driving the Surround Panner plugin[4] in Ableton Live. These LFO-based modulations produce continuously varying channel distribution and diffusion patterns;

(2) The lighting process is driven by the same trigger stream and establishes an event-accurate audiovisual coupling. Trigger events are transmitted as MIDI Note On messages generated in Max to a Wolfmix MKII lighting controller via a USB-MIDI interface, activating randomly selected lighting effect patterns whose activation is terminated by an end-of-playback trigger generated at the end of the audio sample. The end-of-playback trigger recalls a static lighting state, ensuring a deterministic reset of the visual system between events.

Within the Wolfmix MKII, all six LED fixtures are assigned to Group A and are addressed through individual presets to enable discrete spatial control of light sources. A dedicated Preset 0, mapped to MIDI note 60, corresponds to the static lighting state and is automatically recalled at the end of each sample. Presets 1–6 are assigned to MIDI notes 61–66 (velocity 127) through the Wolfmix Mappings section, allowing direct software-level actuation from Max.

This shared trigger-based control architecture allows a precise audiovisual coupling between vocal excitation, sound spatialization, and light modulation, supporting the system to operate as a more integrated audiovisual instrument rather than as loosely synchronized audio and lighting subsystems.

A complete diagram of the system architecture is shown in Fig. 1.

Moreover, each vocal input is processed through the following Ableton Live effects chain: it first passes through a vector delay Max for Live device[5], configured with “Scale” = 1.00, “Speed” = 1.00, and a wet/dry mix value of 23. To introduce continuous parametric variation, two independent LFO modulators are assigned to the “Speed” and “Scale” parameters, both configured to modulate the parameters slowly over time.

Subsequently, the signal is routed through an Audio Effect Rack composed of a resonator filter bank, a Grain Delay module, and a Reverb unit, all native devices in Ableton Live 12. Additional LFO-based modulation is applied within this rack to produce evolving spectral filtering, micro-temporal dispersion, and variable reverberant depth, thereby enhancing the perception of vocal fragmentation and spatial diffusion.

From a spatial perspective, the processed vocal signals are routed primarily to the frontal loudspeaker pair (channels 1–2 of the audio interface). In contrast, the pre-composed background texture is diffused through the rear loudspeakers (channels 3–4).

5 Installation Requirements

For the installation setup, the organizers are only required to provide four loudspeakers for quadraphonic diffusion and an appropriate space to host the installation. All other technical components, including the audio interface, microphones, lighting system, computer, and software environment, will be provided by the authors.

As previously mentioned, the installation was presented in the foyer of the Music Conservatory of Pescara. In response to the specific characteristics and dimensions of the space, the microphones were suspended from the ceiling using hooks,

[3] https://cycling74.com/products/max-9

[4] https://www.ableton.com/en/packs/surround-panner/

[5] https://www.ableton.com/en/packs/inspired-nature/

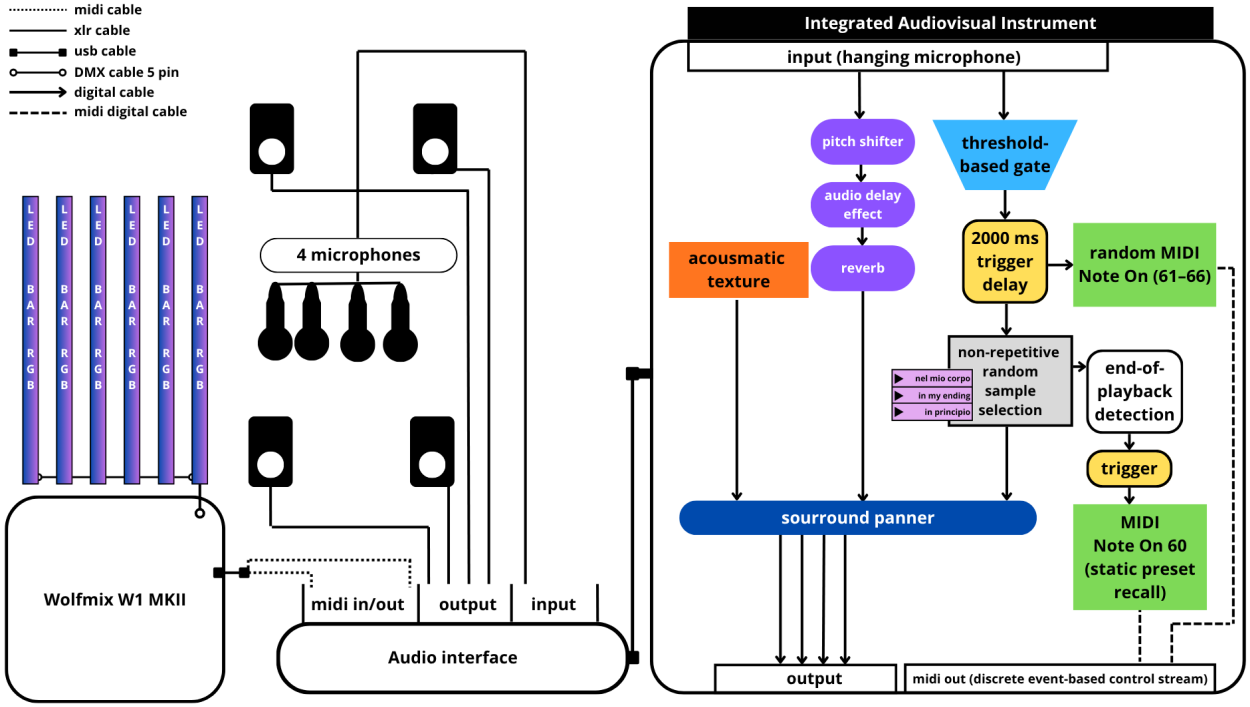


Fig. 1. A diagram that illustrates the hardware and software pipeline of the proposed installation.

without drilling or causing any structural alteration, while the LED bars were strategically positioned according to the spatial configuration. This arrangement can be seen in the attached video in Section 6.

It is important to note that the installation does not necessarily require ceiling-mounted microphones; microphones on regular floor stands are a viable alternative. The installation is fully adaptable and can be configured to suit any space assigned.

6 Media Links

- Video: https://www.youtube.com/watch?v=zQ4OcgfCfdY

7 Ethical Standards

This study followed all ethical and data protection guidelines from the Conservatory of Music of Pescara. This paper complies with the NIME Conference standard [5]. We also declare no conflict of interest.

References

[1] David Birchfield, Kelly Phillips, Assegid Kidané, and David Lorig. 2006. Interactive public sound art: a case study. In *Proceedings of the 2006 conference on New interfaces for musical expression*. 43–48.

[2] Nina Horvath. 2009. The "Theatre of the Ear": Analyzing Berio's Musical Documentary A-Ronne. *Musicological Explorations* 10 (2009), 73–103.

[3] Rébecca Kleinberger, Nikhil Singh, Xiao Xiao, and Akito van Troyer. 2022. Voice at NIME: a Taxonomy of New Interfaces for Vocal Musical Expression. In *NIME 2022*. PubPub.

[4] Zhengyang Ma, Iurii Kuzmin, Duan Ruilei, and Raul Masu. 2024. Pharosphones: interactive audience participation using light. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 480–485.

[5] Fabio Morreale, Nicolas Gold, Cécile Chevalier, and Raul Masu. 2023. NIME Principles & Code of Practice on Ethical Research. https://doi.org/10.5281/zenodo.7545682

[6] Michael Mulshine. 2023. Near and far. In *Installation Proceedings of the International Conference on New Interfaces for Musical Expression*.