Aalborg University
Department of Architecture, Design and Media Technology
Sound and Music Computing

Sonifying Machinery in Theatrical Performances

A Real-Time Sonification System
for Kinetic Theatrical Scenography

Adam Munch Møller
ammo24@student.aau.dk

Master's Thesis — Sound and Music Computing — May 2026

ABSTRACT

Theatrical machinery can generate rich, multi-variable kinematic data. This project develops SOMATRIC (SOnifier of MAchinery in TheaTRICs), a real-time sonification system for the 14 motorised Wahlberg winches deployed in *Isnætter*, a production at Sydhavns Teater adapting Anna Kavan's novel *ICE* [1]. Rather than mapping individual winch positions to individual audio parameters, the system derives collective emergent metrics - landscape index, kinematic entropy, directional consensus, and wave detection - from the relational behaviour of the full winch array, and routes these to a modular DSP effect chain running as a Max for Live plugin in Ableton Live. Mapping decisions are grounded in embodied conceptual metaphor theory and the production's own dramaturgical vocabulary, ensuring that the sonic output participates in the same thematic network as the choreography, lighting, and scenography. A browser-based operator interface allows non-technical production staff to configure mappings at runtime without knowledge of the underlying signal processing. The system was evaluated through a two-phase protocol: preliminary sessions using synthetic motion senders to test the perceptual legibility of the mapping framework, followed by a planned live session with the full winch infrastructure. The production premiered in May 2026; the sonification system was not integrated into the premiere due to access constraints.

1. INTRODUCTION

This report documents the second iteration of an ongoing collaboration with Sydhavns Teater, developed in connection with their production *Isnætter* - an adaptation of Anna Kavan's novel *ICE* [1]. The production deploys 14 motorised Wahlberg winches as scenographic instruments, using them to shape, raise, and collapse large plastic textile surfaces across the stage. These winches transmit real-time positional data over OSC, producing a continuous stream of kinematic information that changes with every movement of the fabric.

Figure 1: The plastic textile surface in *Isnætter* [2], shaped by the winch array above a performer. Photo: Sydhavns Teater.

The expressive range of this winch infrastructure has been demonstrated in other artistic contexts. *Aerial Waves* [3], a spatial installation exhibited at Sydhavnens Festival and Aarhus Festuge, used a grid of twelve kinetic winches to animate a 4×6 m textile ceiling, allowing visitors to shape the space through a kinetic interface. The producer Fred Again... deployed 84 custom Wahlberg winches across a 1,200 m^2 semi-transparent fabric roof installation for the USB002 world tour, spanning ten concerts in ten cities over ten weeks, with creative direction by visual artist Boris Acket [4]. In both cases the winch movement functioned as a purely visual and spatial medium: the control data driving the fabric was not used to produce or shape sound. The present project addresses this absence directly.

Figure 2: The USB002 world tour fabric installation by Wahlberg Motion Design and Boris Acket, deployed across a 1,200 m^2 semi-transparent roof. Image: Wahlberg Motion Design [4].

The central challenge this project addresses is how to translate that data stream into a meaningful sonic contribution to the performance. A naive approach - mapping individual winch positions directly to individual audio parameters - produces output that is sonically active but dramaturgically arbitrary: the sound moves with the machines but communicates nothing about the world the performance is trying to construct. *Isnætter* operates through a dense network of thematic axes - thermal, topological, kinetic, psychological - and the production's scenography, choreography, and lighting are all designed to express these axes simultaneously. For the sound to participate in that network, the mappings must be grounded in the same dramaturgical logic.

SOMATRIC proposes a mapping framework in which the primary sonic parameters are derived not from individual winch states but from the collective, relational behaviour of the full winch array: metrics such as spatial entropy, directional consensus, landscape index, and kinematic variance that are arbitrary for any single winch but emerge from the system as a whole. These collective parameters are computed in real time by a Node.js bridge and routed to a modular DSP effect chain running as a Max for Live plugin in Ableton Live. A browser-based operator interface, developed in response to feedback from the first iteration, allows non-technical production staff to configure and adjust mappings during rehearsal without knowledge of the underlying signal processing.

The system was designed for deployment across the full 14-winch array over the 75-minute arc of the performance. The production premiered in May 2026 without the sonification system integrated, as access to the winch infrastructure during the final rehearsal period was not available. The system remains available for future deployment.

2. RESEARCH QUESTIONS

The storyboard of *Isnætter* organises the performance around a set of recurring thematic qualities - temperature, topology, movement, and psychological coherence - expressed simultaneously through the choreography, the lighting, and the movement of the winch system. Sound is absent from this network. This project asks whether it can participate on the same terms: not as accompaniment, but as a medium that carries the same metaphorical content through a different sensory channel.

This gives rise to the following research question:

> Can convincing sonic metaphors be derived from the data produced by mechanical theatrical set pieces - and can those metaphors be made perceptually legible to a listener without prior instruction?

This question concerns perceptual validity: whether the embodied metaphors identified in the dramaturgical analysis survive translation into the sonic domain, and whether the resulting sound communicates its intended meaning to someone who has not been told what the mappings are.

3. STATE OF THE ART

3.1 Parameter Mapping in Electronic Instruments

The relationship between a performer's physical gestures and the resulting sound in electronic music is mediated by a mapping layer - a set of rules translating input parameters to DSP parameters. Early electronic instruments often employed direct one-to-one mappings, where a single input controls a single output. Hunt and Kirk were among the first to treat this mapping layer as a distinct design concern, separate from both the interface and the sound generator, arguing that the choice of mapping strategy fundamentally defines the character of an instrument [5]. Hunt and Wanderley [6] extended this finding empirically, demonstrating that complex many-to-many mappings - where multiple inputs simultaneously influence multiple outputs - produce richer and more expressive long-term performance than simple one-to-one relationships. Crucially, they frame the mapping not as a neutral technical bridge but as the expressive core of the instrument itself. This provides a precedent and motivation for the present project's use of collective, relational parameters: rather than routing individual winch positions to individual effect parameters, the system derives higher-order quantities from the behaviour of the full array.

3.2 Embodied Metaphor and Sonification Design

A parallel body of literature addresses not just the structure of mappings but their perceptual and cognitive grounding. Lakoff and Johnson [7] established that human conceptual systems are fundamentally metaphorical in nature, and that primary conceptual metaphors arise from recurring correlations between embodied physical experience and abstract domains. Mappings such as WARMTH IS INTIMACY, HEAVINESS IS LOW, and OPENNESS IS FREEDOM are not cultural conventions but pre-linguistic structures rooted in bodily experience. This has direct implications for sonification design: mappings that align with primary conceptual metaphors will feel intuitively meaningful to listeners without requiring explicit instruction. Roddy and Bridges [8] applied this framework directly to the sonification mapping problem, arguing that embodied image schemata and conceptual metaphor theory provide a principled basis for mapping selection. Rather than choosing mappings arbitrarily or purely on technical grounds, designers can identify metaphorically congruent relationships between data dimensions and sonic parameters - relationships that carry intrinsic perceptual logic. Walker and Kramer [9] further demonstrated that the directionality of a mapping matters perceptually, showing that temperature-to-pitch mappings feel natural only when their direction aligns with existing conceptual metaphors. Dubus and Bresin [10], in a systematic review of 179 sonification studies, found that spatial auditory dimensions are almost exclusively used to represent kinematic quantities - movement, velocity, acceleration, trajectory. This finding directly legitimizes the present project's approach of mapping kinetic winch data to spatial audio parameters, situating it within established sonification practice [11].

3.3 Spatial and Topological Sound Description

A third relevant tradition comes from electroacoustic music theory. Smalley [12] developed spectromorphology as a framework for describing sound shapes and motion processes, grounding them explicitly in embodied physical and environmental experience. Notably, Smalley identifies characteristic motion types - including push/drag, flow, rise, and drift - as fundamental categories of sonic behavior, and argues that spatial metaphors such as emptiness/plentitude and diffuseness/concentration are intrinsic to how we perceive and describe sound. His framework establishes that topology and spatial geometry are not merely visual or physical concepts but deeply embedded in sonic perception. Smalley's motion categories bear a striking resemblance to both the choreographic score vocabulary of the performance examined in this project and the algorithmically detected gesture categories developed as part of the mapping system. This convergence across three independent sources - electroacoustic theory, choreographic practice, and computational gesture detection - suggests that these motion categories reflect genuine perceptual and cognitive universals rather than domain-specific conventions.

3.4 The Gap This Project Addresses

The closest existing work is Baalman, Salter, and Moody-Grigsby's *Schwelle* [13], a live theatrical performance driven by real-time wireless sensors - accelerometers on a performer's body and photoelectric sensors distributed across the stage - mapped through a dynamical-systems engine to continuous, evolving sound design. *Schwelle* establishes a precedent for separating the sensing, mapping, and compositional layers in live theatrical sonification, and demonstrates that real-time kinematic data can carry dramaturgically meaningful sonic content in a performance context. The present project shares this architecture but diverges in two respects: the data source is industrial machinery rather than a human performer, and the primary mapping parameters are collective rather than individual - derived from the relational behaviour of the full winch array rather than from any single sensor.

Collectively, the existing literature establishes that mapping complexity matters, that metaphorical congruence improves perceptual meaningfulness, and that spatial and topological metaphors are cognitively grounded in embodied experience. However, the literature consistently addresses mappings from individual parameters to individual outputs. No existing work addresses the sonification of collective emergent parameters derived from distributed kinetic systems - values that are meaningless for any single input but arise only from the relational behavior of multiple inputs simultaneously. Furthermore, kinetic textile performance as a sonification context remains essentially unaddressed. The present project addresses both gaps, proposing and implementing a mapping framework in which the primary sonic parameters are derived not from individual winch positions but from the collective spatial and kinematic behaviour of the winch system as a whole. The theoretical grounding developed across the four bodies of literature above - mapping complexity, metaphorical congruence, spatial sonification, and spectromorphological motion categories - is applied simultaneously, rather than drawn on selectively. This is the contribution the existing literature does not make.

4. METHODOLOGY

This project employs an iterative, problem-based design methodology grounded in the research-through-design tradition [14]. Rather than beginning with a fixed hypothesis to be tested under controlled conditions, it treats the design process itself as the primary site of knowledge production: each decision, its justification, and its observed consequences constitute the research contribution. The methodology unfolds across four phases - dramaturgical analysis, system design and implementation, expert evaluation, and reflective iteration - each building on the findings of the last.

4.1 Phase 1: Dramaturgical Analysis

The first phase consists of a close reading of the performance material available to the project: the storyboard and choreographic scores for *Isnætter*, and the source novel *ICE* by Anna Kavan [1]. This analysis serves as the methodological foundation for all subsequent mapping decisions, grounding them in the specific dramaturgical context of the production rather than in abstract technical criteria or arbitrary parameter assignments.

The dramaturgical analysis proceeds by identifying recurring thematic and spatial qualities in the source material - topology, temperature, movement quality, psychological states - and asking whether these qualities have physically or metaphorically motivated sonic analogs. This process is explicitly informed by embodied conceptual metaphor theory [7], which holds that meaningful mappings between domains arise from pre-linguistic correlations in embodied physical experience. A mapping justified by both physical acoustic reasoning and conceptual metaphor theory is treated as more robust than one justified by either alone.

From this analysis, a gesture vocabulary is derived: a set of named collective winch behaviors, each formally defined as a computation over the full set of winch states and each carrying a dual acoustic and metaphorical justification. These gestures - including wave, chaos, suck, blow, net rise, net fall, and the landscape index - form the primary mapping sources of the system. Their derivation from dramaturgical analysis rather than technical convention is central to the project's argument that relational, emergent parameters produce more artistically meaningful mappings than isolated one-to-one relationships.

4.2 Phase 2: System Design and Implementation

The second phase translates the gesture vocabulary and mapping hypotheses into a working prototype. The system architecture comprises three primary layers: a browser-based mapping interface served via a Node.js local server and displayed within Max using the `jweb` object; a Node.js bridge layer handling OSC ingestion, gesture computation, and WebSocket communication to the UI; and a modular DSP effect chain in Max/MSP receiving computed gesture values as control signals.

The system was designed with two simultaneous constraints in mind. First, it had to be technically functional - reliably computing collective gesture parameters from live OSC data and routing them to appropriate effect parameters in real time. Second, it had to be operationally accessible - usable by non-technical operators at Sydhavns Teater in a live theatrical context without specialist knowledge of the underlying signal processing. These two constraints were treated as equally important, as a system that is technically correct but operationally opaque fails its primary deployment context.

Implementation proceeded iteratively, with gesture detection algorithms and effect mappings refined in parallel as the system was tested against live and simulated winch data. The effect pipeline is organized as a modular chain of encapsulated processors with consistent inlet-outlet contracts, allowing incremental implementation without disrupting the overall routing architecture.

Figure 3 shows the planned project timeline. The colour categories are a template artefact used for visual separation only and do not reflect a formal risk assessment.

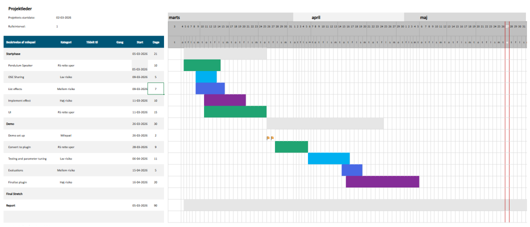

Figure 3: Planned project timeline. Colour categories serve as visual separators only.

The actual process diverged from this plan in several respects. The Pendulum Speaker concept - an early design in which winch motion would drive pendulum-style resonance - was removed before reaching a testing phase; the theatre never confirmed it as a desired element, so development did not proceed beyond the initial planning stage. The plugin conversion pathway was never completed: repeated attempts to establish which audio software Sydhavns Teater would use for the production did not produce a clear answer, making it impossible to build toward a confirmed deployment target; the Max for Live workflow was therefore retained throughout. OSC synchronisation with the actual crane infrastructure was only established very late in the development cycle, meaning the majority of system testing and evaluation used synthetic senders rather than live winch data. The testing and parameter tuning phase, planned to follow the first evaluation round, was partially realised: rather than a full revision pass, parameters were incrementally adjusted between sessions in direct response to evaluator feedback, with threshold and range refinements applied to the mappings that had scored lowest in the preceding session.

4.3 Phase 3: Expert Evaluation

The prototype was evaluated through structured sessions with a combination of expert participants drawn from three relevant domains. First, practitioners from Sydhavns Teater - including members of the production team with direct knowledge of the dramaturgical intentions of *Isnætter* - provided evaluation grounded in the specific performance context. Second, sound designers and practitioners with experience in interactive music systems and live electronic performance provided evaluation from a technical and artistic perspective.

This combination of expert profiles was deliberate. Different domains of expertise surface different dimensions of the system's behavior: production practitioners are best positioned to assess dramaturgical coherence, sound designers to assess sonic quality and mapping expressivity, and academic peers to assess methodological rigor and theoretical grounding. Triangulating across these perspectives produces a more complete evaluation than any single domain could provide.

Each evaluation session addressed two primary dimensions:

- **Ease of use** - Participants interacted with the browser-based interface and were asked to assign mappings, configure effect parameters, and navigate the system for a hypothetical scene. Observations and participant feedback assessed whether the interface affords intuitive operation by non-technical users, and whether the conceptual model of the system - gestures mapped to effects - is legible without prior instruction.

- **Metaphorical coherence** - Participants listened to the system responding to winch data and were asked whether the sonic behavior communicates the intended dramaturgical meanings. This dimension evaluated not technical correctness but perceptual and artistic validity: whether the mappings feel meaningful, whether the sonic character of a wave gesture is distinguishable from a chaos gesture, and whether the hot/cold and mountain/valley axes are perceptually legible in the output.

Evaluation sessions were conducted as semi-structured interviews combined with hands-on interaction with the prototype, allowing participants to articulate both immediate reactions and considered reflections. Sessions were documented through notes and, where consent was given, audio recording.

4.4 Phase 4: Iteration and Reflection

Findings from the expert evaluation sessions informed a second design iteration, in which both the interface design and the mapping strategies were refined in response to observed shortcomings and emergent insights. Ease-of-use findings drove interface revisions - labeling, layout, interaction model. Metaphorical coherence findings drove mapping revisions - effect selection, parameter ranges, gesture threshold tuning.

This cycle of analysis → implementation → evaluation → refinement reflects the problem-based learning approach central to the SMC programme, treating each iteration not as a failure to be corrected but as a structured source of design knowledge. The intended deployment context was the full *Isnætter* production at Sydhavns Teater; access to the winch infrastructure during the final rehearsal period was not available, so Phase 3 was conducted under preliminary conditions only. Observations from the evaluation sessions informed the reflective discussion and conclusions of the report.

5. METAPHORICAL MAPPINGS

The dramaturgical material of *Isnætter* organises itself around four intersecting axes of meaning: a thermal axis (warmth versus cold), a topological axis (mountain versus valley), a kinetic quality axis (stillness through to storm), and a coherence axis (order versus chaos). Each axis is legible in the storyboard both as a choreographic instruction and as a quality of the spatial environment; each also has a well-documented sonic analog grounded in embodied conceptual metaphor theory [7] and supported by empirical sonification literature. The following subsections derive the system's mapping strategy from these axes, citing both the dramaturgical source and the empirical or theoretical justification for each choice.

5.1 Warmth and Cold

The thermal axis is the most persistent quality in both Anna Kavan's source novel and the storyboard of *Isnætter*. Scene 2 is explicitly annotated with the need for warmth (*"der er behov for varme, følelsen af ung kærlighed"*), while the cold encroaches progressively through Dream 1 (Scene 3: *"kulde kommer ind i billedet"*), the Commander's scenes, and culminates in Scene 15 (*"bjerge af is, frostet koldt"*) and Scene 17 (*"de er inde i isen"*). The source of warmth is invariably the Girl as memory (Nanna); the source of cold is invariably the ice, the violence, and the psychotic fantasy.

The metaphor WARMTH IS INTIMACY is identified by Lakoff and Johnson as a primary embodied conceptual metaphor arising from the pre-linguistic correlation of physical warmth with proximity, nurture, and safety [7]. Walker and Kramer demonstrate that mappings feel perceptually natural only when their directionality aligns with the underlying conceptual metaphor: a temperature-to-audio mapping that encodes cold as bright and warm as dark violates the embodied association and produces listener confusion [9]. Dubus and Bresin, in a systematic review of 179 sonification studies, find that temperature representations correlate consistently with spectral brightness and timbral roughness [10]: warm is associated with smooth, low-frequency, intimate timbres; cold with harsh, inharmonic, spectrally diffuse ones.

These findings directly ground the following mappings:

- **Reverb wet/dry → cold.** A fully dry signal is close, present, and warm; a heavily reverberant signal is distant, vast, and cold. The reverb wet parameter is driven by the collective landscape index (mean winch position):

as the fabric rises it opens a larger virtual acoustic space, encoding the approaching ice as an expanding void.

- **Ring modulator carrier frequency → frozen timbre.** Ring modulation destroys the harmonic series and replaces it with sum-and-difference sidebands, producing the metallic, inorganic quality perceptually associated with ice and metal. A carrier near 440 Hz on acoustic material produces a strongly bell-like, frozen colouration; as the carrier frequency rises the timbre becomes increasingly harsh and alien. This directly implements the embodied cold-as-inharmonic association documented by Dubus and Bresin [10].

- **Bitcrusher depth → cold degradation.** Digital quantization noise and sample-rate reduction are perceptually coded as harsh, mechanical, and non-organic - the inverse of the smooth, continuous texture of warmth. The bitcrusher parameters are driven by the kinetic chaos metric (see subsection 5.4), so that physical disorder accumulates as sonic degradation.

5.2 Topology: Mountain and Valley

The topological axis is made explicit by the scenography. Scene 10b names specific cranes (*"Thomas bygger bjerge - koreografi mellem upstage right kraner (7, 10, 11, 13, 14)"*), using the winch system directly as landscape-forming instruments. Scene 15 describes *"bjerge af is"* (mountains of ice), and the closing scenes enclose the stage inside rising plastic walls that constitute the final advance of the ice. The spatial layout of the winch array is thus the primary visual representation of the novel's landscapes.

Smalley's spectromorphological framework identifies emptiness/plenitude and diffuseness/concentration as intrinsic spatial dimensions of sonic perception, arguing that topology is not merely visual but embedded in how listeners perceive and describe sound [12]. Roddy and Bridges apply the embodied image schemata UP IS MORE, CONTAINMENT, and SCALE to sonification design, providing a principled basis for mapping physical height to pitch and spatial extent to acoustic space [8].

- **Mean position (landscape index) → reverb time.** A lowered fabric (valley configuration) maps to a long reverb tail, encoding the enclosed, echoing acoustic of canyon walls and cave interiors. A raised fabric (mountain) maps to a shorter, drier tail - the exposed, open-air quality of a bare peak with no surfaces to reflect.

- **Spatial position variance → quadraphonic width.** When all winches are at similar heights the landscape is flat and the sound field is narrow and centred. When heights are varied - mountains alongside valleys - the spatial position variance is high and the quadraphonic panning field expands, encoding topological complexity as spatial breadth.

- **Upstage-right group average → pitch shift.** The cranes explicitly identified as mountain-builders in Scene 10b belong to a named group whose mean position drives a pitch shift upward. Higher physical position produces higher sonic pitch: the most primitive and universally documented embodied metaphor (UP IS HIGH in pitch) [7], requiring no instruction to be perceptually legible.

5.3 Movement Quality: From Stillness to Storm

The storyboard maps the full duration of the performance onto a kinetic arc from total stillness to catastrophic storm. Scene 0 opens in silence with finely combed sand. Scene 1 is annotated *"rummet begynder at trække vejret"* - the room begins to breathe. Scene 8 describes a chase. Scene 13 activates a plastic wave. The closing storm (around 72 minutes) is described as pure music and scenography: escalating movement ending with the plastic collapsing onto the sand.

Smalley identifies breathing, flowing, pushing, dragging, and turbulence as fundamental spectromorphological motion categories [12], noting that these are grounded in embodied physical experience rather than cultural convention. These categories map directly onto the choreographic score vocabulary of the performance, suggesting that the same underlying perceptual categories are operative in both the kinesthetic and sonic domains. Dubus and Bresin confirm that velocity and acceleration are the kinematic quantities most consistently used to drive spatial audio parameters in sonification practice [10].

- **Stillness (all speeds ≈ 0) $\rightarrow$ near-silence or minimal dry reverb tail.** Rest in the mechanics encodes as rest in the sound. The neutral state policy (subsection 8.3) ensures that unused effects do not accumulate stale values.

- **Breathing (slow periodic collective oscillation) $\rightarrow$ tremolo rate ≈ 0.2–0.5 Hz.** The resting respiratory rate of a human adult is 0.2–0.3 Hz. Mapping slow collective winch oscillation to a tremolo in this range produces a direct cross-modal breathing analog: the room *is* breathing through the sound. This instantiates the primary metaphor RHYTHM IS LIFE [7]: the most universally legible rhythmic pattern in human experience, requiring no instruction to be felt.

- **Directional wave (spatially coherent travelling motion) $\rightarrow$ delay time and stereo panning.** A left-to-right wave produces a stereo ping-pong delay whose period matches the wave crossing time. The directional consensus of the motion vector drives panning direction: sound appears to travel with the wave.

- **Chase (high sustained collective speed) $\rightarrow$ granular density and upward pitch.** Speed drives grain density (denser, more urgent texture) and a slight upward pitch shift. The SPEED IS TENSION association documented by Roddy and Bridges [8] legitimises the pitch increase as a metaphorically motivated choice; Effenberg and Schmitz [15] further demonstrate that pitch-coded kinematic sonification actively modulates listeners' perceived motion speed - sometimes overriding competing visual information - providing perceptual-science grounding for the same association.

- **Storm (maximum speed and maximum chaos simultaneously) $\rightarrow$ full simultaneous effect depth.** Reverb, tremolo, granular, and ring modulation are all driven toward maximum simultaneously. The accumulation of effects encodes overwhelming sensory overload - the sonic correlate of a storm that cannot be resolved into any single quality. Talmy's COMPULSION FORCE schema [16] characterises this escalation: the sonic output is not chosen but compelled by the data.

5.3.0.1 Cross-modal flickering.

Scenes 1 and 3 share a recurring description of visual flickering (*flimren*) in the lighting, and the sound annotation for Scene 1 explicitly instructs that sound should *"lægge sig op af flimren i lyset"* (align with the flickering of the light). Spence's tutorial review of cross-modal correspondences documents a robust systematic relationship between visual flicker rate and auditory amplitude modulation rate: perceivers reliably match flickering light to tremoloed sound at corresponding frequencies [17]. A tremolo rate of 4–10 Hz matches the perceptual frequency range of light flickering and directly implements the cross-modal alignment called for by the storyboard, without requiring the audience to be instructed in the mapping.

5.4 Order and Chaos

The coherence axis runs from the orderly opening (*"sand på gulvet - vel friseret"*, sand on the floor, well-groomed) to the full psychotic fragmentation of Scene 14 (*"tiltagende kaos af sanddynger"*, increasing chaos of sand dunes), where the storyboard notes that the fragmented quality of the choreography is deliberately amplified and the performer loses control of their own violence.

Hunt and Wanderley demonstrate that complex many-to-many mappings - where multiple simultaneous inputs influence multiple outputs - produce richer and more expressive performance than isolated one-to-one relationships [6]. The chaos axis is the system's primary instance of such a mapping: no single winch signals chaos, but the relational behaviour of the array as a whole does.

The system computes a collective kinematic entropy metric defined as the normalised Shannon entropy of the discretised speed distribution across all winches:

$$H(t) = -\frac{1}{\log N} \sum_{k=1}^{N} \hat{p}_k \log \hat{p}_k,$$
$$\hat{p}_k = \frac{|\{i : v_i^{\text{abs}} \in \text{bin}_k\}|}{N}, \qquad (1)$$

where N is the number of winches and $\hat{p}_k$ is the proportion of winches in speed bin k. When all winches move at the same speed $H = 0$ (perfectly ordered); when speeds are uniformly distributed across all bins $H = 1$ (maximally disordered). Its value is a property of the array's collective state, not of any single machine - precisely the class of parameter identified as absent in subsection 3.4.

- **Kinematic entropy → granular jitter.** Low entropy (coherent wave) drives jitter to zero: grains are drawn from a fixed playhead, producing clean loop-based resynthesis. High entropy (chaotic, independent motion) drives jitter to maximum: grains are drawn from random positions in the buffer, producing the scattered, cloud-like texture that Smalley identifies as the sonic analog of turbulence [12].
- **Collective acceleration spread → tremolo depth.** When acceleration magnitudes are spread widely across winches (some jerking, some still), tremolo depth increases, making the amplitude envelope unstable and unpredictable.
- **Kinematic entropy → bitcrusher bit-depth.** Ordered motion preserves full signal fidelity; increasing disorder degrades bit resolution, accumulating the harsh quantization noise that encodes the psychotic fragmentation described in Scene 14. The combined effect of high jitter and low bit-depth produces a sonic degradation that mirrors the collapse of narrative coherence in the final third of the performance. The named metaphor DISINTEGRATION IS FRAGMENTATION captures this directly: coherent motion produces coherent texture; fragmenting motion produces fragmenting sound.

5.4.0.1 Fragmentation and the Girl becoming real.

Anastasia is described in Scene 1 as a *"splintret figur"* (splintered figure), a ghost who can jump in time, while Nanna represents the real, embodied memory. Wishart identifies the transformation from continuous to granular texture as a sonic metaphor for the fragmentation and dissolution of identity [18]. The two representations of the Girl are distinguished sonically by two parameters acting in combination. **Granular coherence** encodes spatial fragmentation: high jitter and low density for Anastasia's scenes (scattered, ghost-like) versus low jitter and high density for Nanna's scenes (coherent, warm, present). **Delay feedback** encodes temporal fragmentation: Anastasia is explicitly described as a figure who can jump in time, existing non-linearly across the performance duration. A stereo delay with elevated feedback driven by the kinematic entropy of the upstage group produces temporal doubling - the same sonic event echoing at a slight remove - encoding the quality of a presence that is never quite anchored in the present moment. This implements MEMORY IS ECHO [19]: past events encoded as sonic recurrences that linger rather than resolve. For Nanna's scenes the delay feedback is reduced to near zero, anchoring the sound in a single unambiguous now. Anastasia's dramaturgical arc - from splintered ghost to a figure who finally sees through the staging - is encoded as a progressive reduction in both jitter and delay feedback over the performance duration, as the system responds to the increasing coherence of collective winch motion in the final scenes. This arc instantiates CONSCIOUS IS UP [20]: as Anastasia's coherence grows, the texture rises from scattered and echoing toward clear and present.

5.5 Summary of Mappings

Table 1 summarises the complete mapping framework, relating each dramaturgical axis to its com-

Axis	Source metric	Effect param.	Grounding
Warmth and Cold			
Warmth	Landscape index ↓	Reverb wet/dry	WARMTH IS PROXIMITY [7]
Cold	Landscape index ↑	Ring mod carrier	Inharmonic = frozen [10]
Cold	Kinematic entropy ↑	Bitcrush depth	Degradation = harsh [10]
Topology: Mountain and Valley			
Mountain	Group avg (up-R) ↑	Pitch shift ↑	UP IS HIGH [7]
Valley	Landscape index ↓	Reverb time long	Enclosed acoustics
Topology	Position variance	QuadPan width	Varied = broad [12]
Movement Quality: Stillness to Storm			
Stillness	All speeds ≈ 0	Effects at neutral	Rest = sonic rest [10]
Breathing	Slow oscillation	Tremolo $\approx 0.3\,\text{Hz}$	Cross-modal breathing [17]
Wave	Directional consensus	Delay/pan direction	Motion in space
Chase	High collective speed	Granular density, pitch	SPEED IS TENSION [8]
Storm	Speed + entropy max	All depths at max	Sensory accumulation
Flicker	Temporal (light cue)	Tremolo 4–10 Hz	Cross-modal rate [17]
Order and Chaos			
Order	Entropy $\rightarrow 0$	Granular jitter $\rightarrow 0$	Coherence = smooth [12]
Chaos	Entropy $\rightarrow 1$	Jitter, bitcrush	Turbulence = scatter [18]
Fragmentation	Scene-specific	Jitter high / density low	Splintered = granular [18]
Temporal split	Entropy (upstage)	Delay feedback ↑	Time-jump = echo [18]

Table 1: Summary of dramaturgically motivated mappings. Source metrics are derived from the collective winch state vector; effect parameters are DSP targets within the modular chain described in section 8.

puted source metric and target DSP parameter.

6. SYSTEM OVERVIEW

The system is structured as a distributed, layered architecture in which raw positional data from the Wahlberg winches is received, semantically interpreted, and translated into audio-effect control signals in real time. Four concerns are separated into distinct layers: data ingestion, semantic computation, operator configuration, and DSP execution. This separation ensures that each layer can be modified or extended independently without disrupting the others - a practical requirement in a live theatrical context where reliability and partial-failure tolerance are essential.

Figure 4 illustrates the full signal flow of the system.

6.1 Data Ingestion Layer

The Wahlberg winches transmit normalised positional data in the range $[0, 1]$ as OSC messages over a local network. Max/MSP receives these messages via a `udpreceive` object listening on port 9000 and immediately forwards the raw values to the Node.js bridge via the `node.script` object. No processing or interpretation occurs at this stage - Max acts purely as a transport receiver, preserving the raw positional stream for semantic interpretation downstream.

6.2 Semantic Computation Layer

The Node.js bridge is the computational core of the system. It maintains an internal model of the winch array, including the spatial positions of each winch as configured by the operator in the browser UI. From this model, Node.js continuously computes a set of collective gesture values - the collective gesture values introduced in section 5: the landscape index, net directional velocity, velocity spread, directional consensus, and wave detection.

The separation of this computation into Node.js rather than Max is a deliberate architectural decision. Node.js affords structured data modelling,

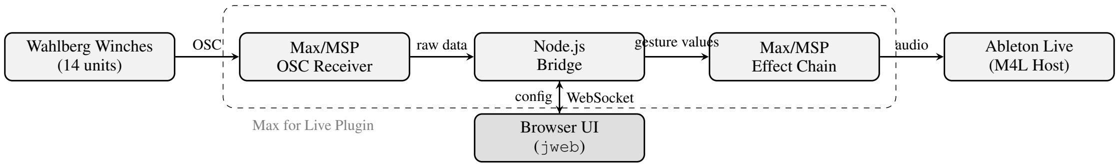

Figure 4: System signal flow. Raw OSC data from the winches enters Max, is passed to Node.js for gesture computation, and returns to Max as computed control values driving the effect chain. The browser UI communicates bidirectionally with Node.js over WebSocket, allowing the operator to configure winch groupings, spatial areas, and effect assignments at runtime.

flexible grouping logic, and clean WebSocket integration with the browser UI - concerns that are cumbersome to express in Max's dataflow paradigm. Max, conversely, handles audio-rate signal processing with lower latency and tighter DSP integration. Each environment is used for what it does best.

6.3 Operator Configuration Layer

The browser-based UI is served locally by Node.js and displayed within Max using the `jweb` object, requiring no external browser or network access beyond the local machine. The UI allows the operator to:

- Position winches on a 2D spatial canvas by dragging, reflecting their physical arrangement in the performance space
- Define winch groups and spatial regions (edge, center, left, right) that inform gesture computation
- Assign effects to gesture parameters and select mapping types
- Store and recall presets for different scenes or dramaturgical states
- Toggle the isometric 3D view off to reduce rendering load
- Hide individual interface panels to reduce visual clutter during performance

Configuration changes are transmitted from the UI to Node.js over WebSocket in real time, immediately affecting how incoming winch data is interpreted. This means the operator can reconfigure the semantic layer of the system without interrupting audio, supporting the dynamic demands of a live theatrical performance.

The UI is designed to be operable by non-technical production staff without knowledge of the underlying signal processing. Mapping assignments are expressed in dramaturgical terms - gesture names and effect labels - rather than technical parameter values, lowering the threshold for meaningful operation in a performance context.

6.4 DSP Execution Layer

Computed gesture values are sent from Node.js back into Max via the `node.script` object, where they are routed to a modular effect chain. Each effect module receives a consistent control-plus-stereo-audio inlet contract, allowing modules to be added, removed, or bypassed without rewiring the overall graph. The effect chain runs as a Max for Live plugin hosted in Ableton Live, which provides the audio interface, session management, and monitoring environment for the performance.

The separation of DSP execution from gesture computation means that the effect chain receives only semantically meaningful control values - not raw winch positions. An effect module responding to the landscape index, for instance, has no knowledge of individual winch heights; it receives only the single derived value representing the collective topological state of the array. This encapsulation keeps each layer focused and maintainable, and ensures that the artistic logic of the mappings is expressed in the computation layer rather than distributed across the DSP graph.

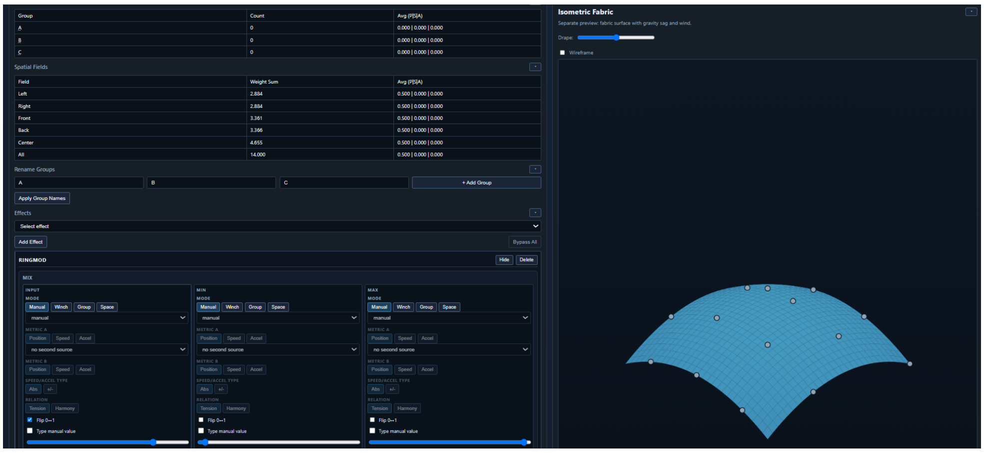

Figure 5: Browser UI: isometric view of the winch array, providing a naive topological representation of the fabric surface based on winch positions. The view does not account for draping; it serves as a spatial orientation aid showing the relative height configuration of the 14 attachment points.

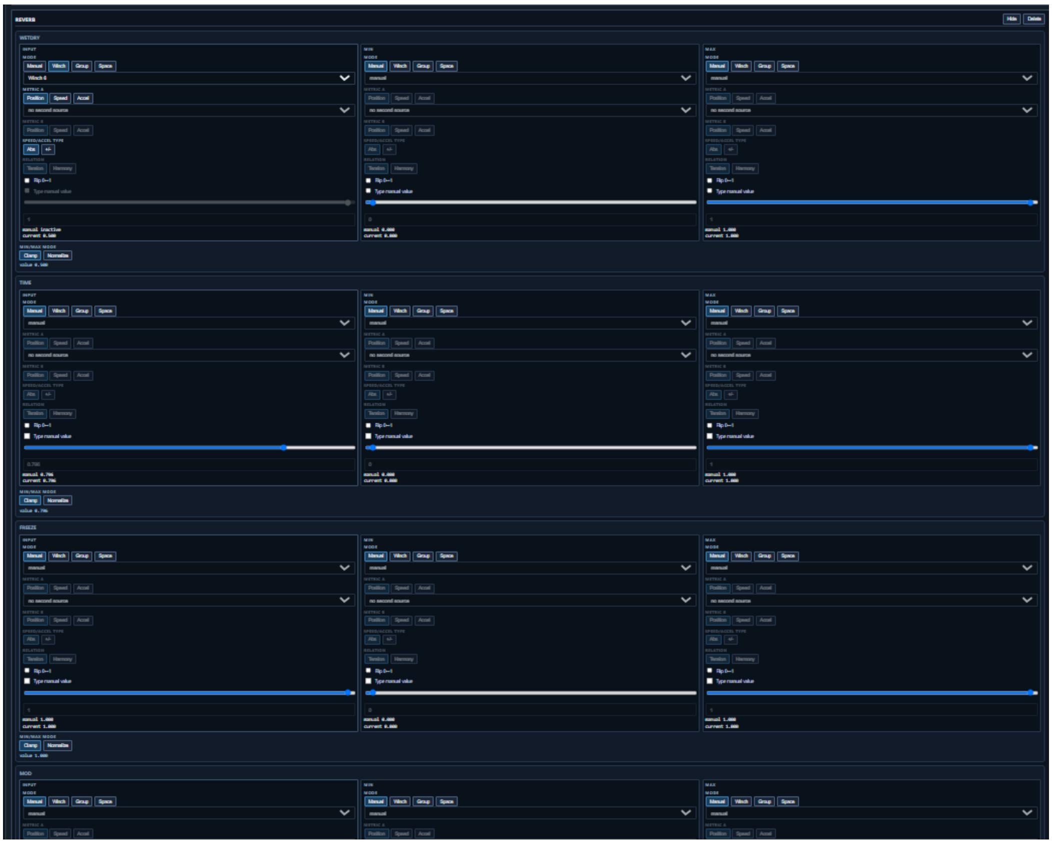

Figure 6: Browser UI: effect assignment panel. Each gesture parameter can be routed to one or more DSP effects; the operator sets mapping type and min/max range.

7. IMPLEMENTATION

The system was implemented as a real-time control and audio-processing pipeline spanning four layers: a Node.js bridge, a Max/MSP DSP patch, OSC-compatible control transport, and Ableton Live integration through Max for Live. The design goal was to keep effect control deterministic while allowing external position and mapping data to drive parameters at runtime.

7.1 Node_js

The Node.js layer acts as a bidirectional bridge between browser UI clients and Max/MSP. It provides:

- An HTTP server for the UI (`localhost:8080`)
- A WebSocket server for low-latency control messages
- Message translation from JSON payloads to Max outlet commands
- Broadcast of Max-originated data back to connected UI clients

Incoming UI messages include `effect_values`, `effect_mapping`, `winch_data`, and configuration summaries. Outgoing bridge messages are normalized into Max-friendly aliases (e.g. `from_browser_volume_gain`, `from_browser_delay_time`) to avoid fragile string parsing in the DSP patch.

The bridge also supports compatibility routing for indexed position streams (`/pos1 ... /pos14`) and list-based winch updates. This allowed different sender formats to converge into a common internal representation without rewriting the DSP layer.

7.2 Max

The Max patch implements the real-time audio chain and effect modules. Audio enters via `plugin~` and exits via `plugout~`, enabling direct use as an audio effect context.

The effect chain is modular, with each module receiving:

1. a control inlet (browser/bridge commands),
2. left and right audio signals,
3. and producing processed stereo output.

Implemented modules include:

- **Volume**: gain scaling
- **Pitch**: `pitchshift~`-based transposition
- **Delay**: feedback delay network
- **Reverb**: built-in `reverb~`-based wet path with wet/dry control
- **Pan/QuadPan**: stereo/spatial balance controls
- **Filter**: `lores~`-based tone shaping with cutoff, resonance, and dry/wet blend

To improve runtime robustness, control parsing was hardened so modules can receive consistent alias-style commands, and neutral parameter resets are used when effects are removed.

7.3 OSC

OSC support was implemented through Max UDP receive and bridge normalization. Two practical OSC patterns were considered:

- multi-value payloads in one message,
- one-float-per-address streams (e.g. `/posX`).

The implemented path accepts position updates and maps them into the internal winch state vector:

$$\mathbf{p}(t) = [p_1(t), p_2(t), \ldots, p_{14}(t)], \quad p_i \in [0, 1].$$

From this state, derived control metrics (including smoothed motion features) can be used by effect mappings in the UI. This separation ensures OSC schema differences are handled at the transport edge, while DSP control remains unchanged.

7.4 Ableton

Integration with Ableton Live was achieved through a Max for Live audio-effect workflow:

1. Build/load the device as an `.amxd` Audio Effect
2. Place the device on an Ableton audio track

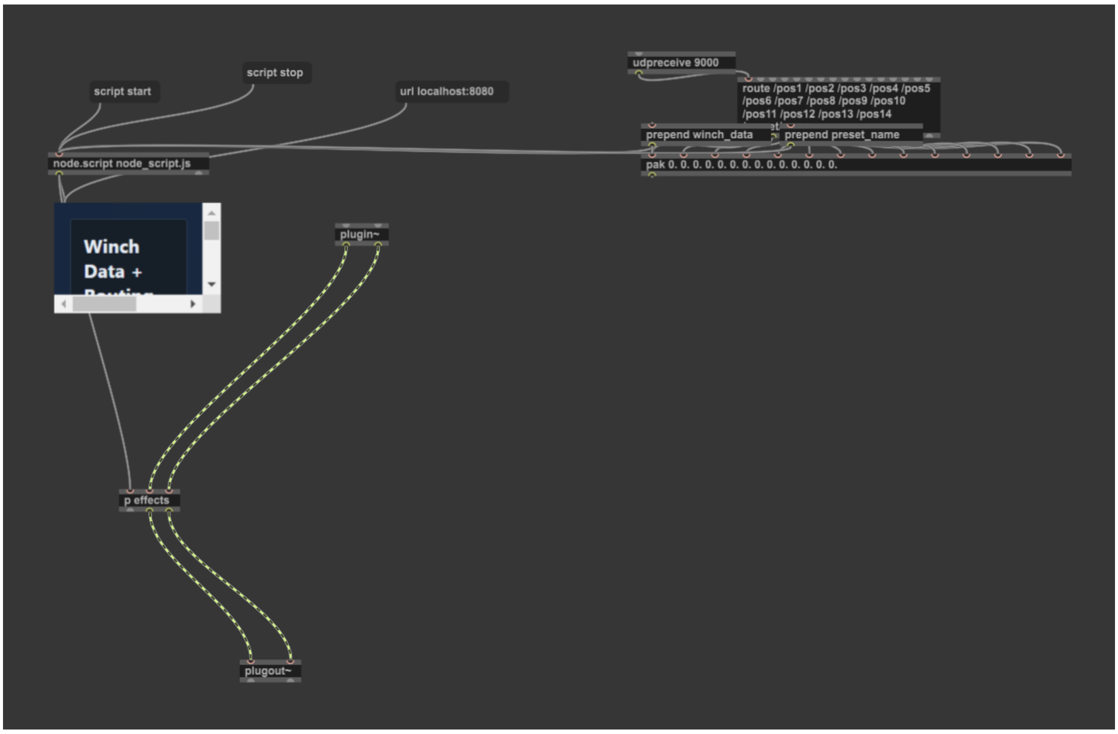

Figure 7: Max/MSP outer patch: the top-level signal flow, showing the `plugin~`/`plugout~` audio interface, the `node.script` bridge object, and the modular effect chain routing.

3. Use track audio as the live source into the Max DSP chain
4. Control effects through the embedded UI/bridge path

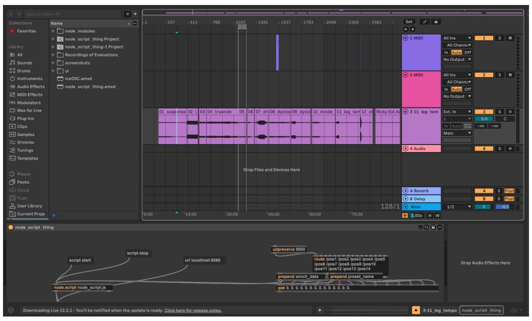

Figure 8: The Max for Live device as it appears in Ableton Live, with the browser UI rendered inside the `jweb` panel and the effect chain running as an audio effect on the production's soundtrack track.

This avoided external virtual-audio routing and kept the audio path native to Live. During testing, a reliable workflow was to:

- validate dry passthrough first,
- enable one effect at a time,
- apply extreme parameter values for audible verification,
- and confirm neutral reset after effect deletion.

The resulting implementation supports practical live testing and iterative refinement, while keeping control transport, mapping logic, and DSP processing cleanly separated.

8. EFFECT PROCESSING PIPELINE

The system translates raw positional data from motorised winches into real-time audio-effect parameters through a cascade of computation stages:

kinematic state estimation, spatial region weighting, responsive parameter mapping, and modular DSP execution. Each stage is described formally below.

8.1 Kinematic State Estimation

Let $p_i(t) \in [0,1]$ be the normalised position of winch i at time t, where $p = 0$ corresponds to fully lowered and $p = 1$ to fully raised. On each incoming data frame, the elapsed time is computed as

$$\Delta t = \max\left(\frac{t - t_{\text{prev}}}{1000},\ 0.016\right)\ \text{s},$$

floored at one nominal frame (16 ms) to prevent division-by-zero during burst arrivals. The raw signed velocity is then

$$\dot{p}_i(t) = \frac{p_i(t) - p_i(t_{\text{prev}})}{\Delta t}.$$

Because raw finite differences are noisy, an exponential moving average (EMA) with smoothing coefficient $\alpha = 0.25$ is applied to obtain a smoothed velocity estimate:

$$\hat{v}_i(t) = \hat{v}_i(t_{\text{prev}}) + \alpha\big[\dot{p}_i(t) - \hat{v}_i(t_{\text{prev}})\big].$$

Acceleration is derived from the change in smoothed velocity and then independently smoothed with the same α:

$$\ddot{p}_i(t) = \frac{\hat{v}_i(t) - \hat{v}_i(t_{\text{prev}})}{\Delta t},$$
$$\hat{a}_i(t) = \hat{a}_i(t_{\text{prev}}) + \alpha\big[\ddot{p}_i(t) - \hat{a}_i(t_{\text{prev}})\big]. \quad (2)$$

All four kinematic signals are then normalised into the $[0,1]$ range expected by the mapping layer. Throughout the following, $\text{clamp}_{[a,b]}(x) = \max\big(a,\ \min(b,\ x)\big)$ denotes the saturation function that constrains x to the interval $[a,b]$. Absolute values represent magnitude (useful for detecting activity regardless of direction); signed values are centre-encoded around 0.5 (useful for directional mappings):

$$v_i^{\text{abs}} = \text{clamp}_{[0,1]}\left(\frac{|\dot{p}_i|}{\nu_v}\right), \quad (3)$$

$$v_i^{\text{sgn}} = \text{clamp}_{[0,1]}\left(\frac{1}{2} + \frac{\dot{p}_i}{2\nu_v}\right), \quad (4)$$

$$a_i^{\text{abs}} = \text{clamp}_{[0,1]}\left(\frac{|\hat{a}_i|}{\nu_a}\right), \quad (5)$$

$$a_i^{\text{sgn}} = \text{clamp}_{[0,1]}\left(\frac{1}{2} + \frac{\hat{a}_i}{2\nu_a}\right), \quad (6)$$

where $\nu_v = 2$ (units s^{-1}) and $\nu_a = 6$ (units s^{-2}) are normalization constants tuned empirically against live winch data to avoid sustained saturation under typical performance motion.

8.2 Spatial Region Weighting

Five named spatial regions - Left, Right, Front, Back, and Center - are defined as soft Gaussian-like influence zones. Each region r has a reference centre $\mathbf{c}_r$ in the normalised canvas coordinate space $[0,1]^2$:

$$\mathbf{c}_{\text{left}} = (0,\ 0.5), \quad \mathbf{c}_{\text{right}} = (1,\ 0.5),$$
$$\mathbf{c}_{\text{front}} = (0.5,\ 1), \quad \mathbf{c}_{\text{back}} = (0.5,\ 0),$$
$$\mathbf{c}_{\text{center}} = (0.5,\ 0.5).$$

The spatial weight of winch i relative to region r is a linear falloff from unity at the centre to zero at the diagonal corner of the unit square:

$$w_{ir} = 1 - \text{clamp}_{[0,1]}\left(\frac{\|\mathbf{x}_i - \mathbf{c}_r\|}{d_{\max}}\right),$$
$$d_{\max} = \tfrac{1}{\sqrt{2}} \approx 0.707, \quad (7)$$

where $\mathbf{x}_i = (x_i, y_i)$ is the canvas position assigned to winch i by the operator. The weighted regional average of metric $m \in \{p, v^{\text{abs}}, v^{\text{sgn}}, a^{\text{abs}}, a^{\text{sgn}}\}$ is

$$\bar{m}_r = \frac{\sum_i w_{ir}\, m_i}{\sum_i w_{ir}}.$$

This formulation allows winches at the physical edges to contribute strongly to "Left" or "Right" while central winches contribute proportionally to all regions, producing smooth, artefact-free transitions as the fabric configuration changes.

8.3 Parameter Mapping

Each effect parameter is driven by a *source control* object that specifies a mapping mode, a data source, optional bounds, and a modulation depth $d \in [0, 1]$. The pipeline supports four mapping modes:

Manual. The parameter is fixed at a constant value set by the operator; no live data is used.

Winch. The source is a single winch's kinematic metric $m_i \in \{p_i, v_i^{\text{abs}}, \ldots\}$.

Group. The source is the unweighted mean of a named winch group: $\bar{m}_G = \frac{1}{|G|} \sum_{i \in G} m_i$.

Space. The source is a soft-weighted spatial regional average $\bar{m}_r$ as defined above.

For Group and Space modes, a second source may optionally be specified to compute a *relational* value between two regions A and B:

$$\begin{aligned} \text{tension}(A, B) &= |\bar{m}_A - \bar{m}_B|, \\ \text{harmony}(A, B) &= 1 - |\bar{m}_A - \bar{m}_B|. \end{aligned} \tag{8}$$

Tension measures spatial or kinematic divergence between regions; harmony measures convergence. Both are in $[0, 1]$.

8.3.0.1 Depth scaling.

Given source value $s \in [0, 1]$, depth modulation is applied before bounds clamping. For unipolar sources (position, absolute speed/acceleration):

$$s_{\text{d}} = \text{clamp}_{[0,1]}(s \cdot d).$$

For signed motion inputs, the value is treated as a bipolar modulator centred on 0.5 (corresponding to zero motion):

$$s_{\text{d}} = \text{clamp}_{[0,1]}\left(\frac{1}{2} + \frac{(2s - 1)\, d}{2}\right).$$

8.3.0.2 Output range and mode.

Each parameter additionally specifies a dynamic lower bound l and upper bound h, each of which is itself a source control object. Two output scaling modes are available:

- **Clamp:** $v_k = \text{clamp}(s_{\text{d}},\, l,\, h)$.
- **Normalize:** $v_k = \text{clamp}_{[0,1]}\left(\frac{s_{\text{d}} - l}{h - l}\right)$, where the depth-modulated signal is first remapped to fill the full $[0, 1]$ output range defined by $[l, h]$.

8.3.0.3 Neutral state policy.

When an effect is inactive or absent from the operator's current configuration, all of its parameters are explicitly driven to predefined neutral values v_k° each computation cycle. Representative neutrals include $v^\circ_{\text{volume.gain}} = 1$ (unity gain), $v^\circ_{\text{reverb.wetDry}} = 0$ (fully dry), and $v^\circ_{\text{pitch.mix}} = 0$ (pitch effect fully bypassed). This prevents stale or uninitialised values from persisting in the DSP graph when effect modules are added, removed, or reconfigured mid-performance.

8.4 Effect Modules

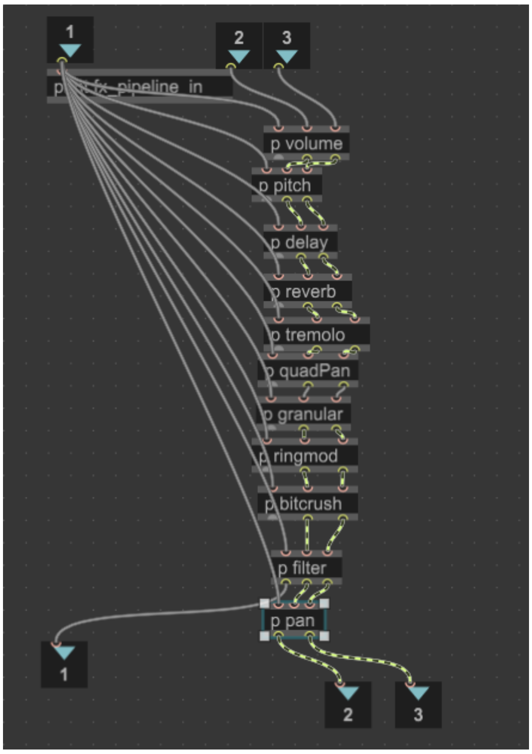

Figure 9: Max/MSP inner patch showing the full serial effect pipeline. Each module receives a stereo audio pair and a control inlet from the Node.js bridge, and passes processed audio to the next stage.

Eleven DSP modules are implemented as encapsulated Max sub-patchers connected in a fixed serial chain:

$$\text{volume} \to \text{pitch} \to \text{delay} \to \text{filter} \\ \to \text{reverb} \to \text{tremolo} \to \text{quadPan} \\ \to \text{granular} \to \text{ringmod} \to \text{bitcrush} \to \text{pan}.$$

Each module receives a control inlet carrying parameter messages from the Node.js bridge and a stereo pair of audio-rate signal inlets, and produces a processed stereo output. Modules not added by the operator pass audio through unchanged; no silence or artefact is introduced by an unused slot.

8.4.0.1 Perceptual palette.

Each module contributes a distinct sonic quality to the chain. **Reverb** encodes spatial distance: dry is close and present; increasing wet depth opens the sound into a larger, colder acoustic space. **Pitch shift** encodes altitude: upward transposition carries the instinctive weight of height and tension; downward, of gravity and descent. **Delay** encodes temporal depth: low feedback doubles and thickens; high feedback accumulates haunting repetitions. **Filter** colours the spectrum: high-pass thins toward cold clarity; low-pass darkens toward warmth; band-pass isolates a resonant region. **Tremolo** encodes rhythmic life: at 0.2–0.3 Hz it breathes; at higher rates it flickers and agitates. **Bitcrusher** introduces digital degradation - coarse quantization noise and sample-rate reduction - producing a harsh, mechanical texture that reads as non-organic. **Ring modulation** destroys the harmonic series, replacing it with metallic or bell-like sidebands depending on the carrier frequency. **Pan and quadPan** place sounds in physical space, encoding the spatial tilt and distribution of the fabric as auditory position in the room.

8.4.1 Volume

The volume module applies a single gain coefficient g to both channels:

$$y_L(t) = g \cdot x_L(t), \qquad y_R(t) = g \cdot x_R(t), \\ g \in [0, \infty).$$

The default $g^\circ = 1$ preserves unity gain. In practice the parameter range is bounded to $[0, 1]$ by the mapping layer unless the operator explicitly overrides it.

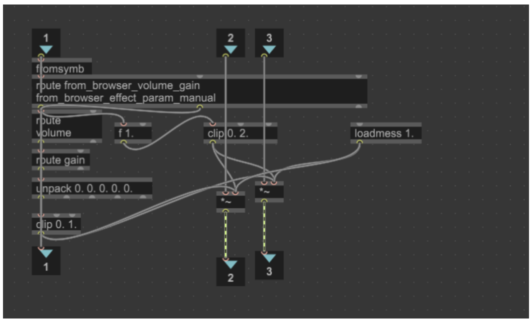

Figure 10: Max/MSP volume module sub-patcher.

8.4.2 Pitch Shift

Pitch transposition is implemented using Max's `pitchshift~` object. The SHIFT parameter $s \in [0, 1]$ is centre-encoded: $s = 0.5$ corresponds to unity transposition, $s = 0$ to a full downward shift, and $s = 1$ to a full upward shift (typically ± 12 semitones). The module operates as a wet/dry blend:

$$y(t) = m \cdot \text{pitchshift}(x(t), s) + (1 - m) \cdot x(t),$$

where $m \in [0, 1]$ is the mix parameter (neutral: $m^\circ = 0$, i.e., fully dry by default).

8.4.3 Delay

The delay module implements a stereo feedback delay line. Three parameters govern its behaviour:

- **time** $\tau \in [0, 1]$: delay length, linearly scaled to an absolute time range in milliseconds.
- **feedback** $f \in [0, 1]$: fraction of the delayed signal fed back into the delay input. The recirculated signal decays exponentially with each tap: $\text{level}_n = f^n \cdot \text{level}_0$.
- **mix** $m \in [0, 1]$: wet/dry blend (neutral: $m^\circ = 0$, fully dry).

Defaults ($\tau^\circ = 0.2$, $f^\circ = 0.4$) produce a moderate echo that is inaudible when $m = 0$ and becomes progressively present as mix is raised.

8.4.4 Reverb

The reverb module wraps Max's built-in `reverb~` object with a four-parameter interface:

- **wetDry** $w \in [0, 1]$: linear crossfade between dry and fully wet signal (neutral $w^\circ = 0$, fully dry).
- **time** $\tau \in [0, 1]$: reverb tail decay (default $\tau^\circ = 0.35$, corresponding to a medium hall).
- **freeze** $z \in \{0, 1\}$: when active, the reverb tail is held indefinitely, producing a sustained spectral pad from any transient input [21].
- **mod** $\mu \in [0, 1]$: pitch modulation depth within the reverb tail, useful for producing slow shimmer effects [21].

8.4.5 Filter

The filter module implements a parallel three-mode topology: a shared second-order filter stage produces low-pass, band-pass, and high-pass outputs simultaneously, and the operator controls the contribution of each mode independently. The output signal is:

$$y(t) = m \cdot \big(\ell_{\text{LP}} \cdot \text{LP}(x) + \ell_{\text{BP}} \cdot \text{BP}(x) + \ell_{\text{HP}} \cdot \text{HP}(x)\big) + (1 - m) \cdot x(t),$$

where LP, BP, and HP share a common cutoff frequency and resonance. The four operator-facing parameters are:

- **low** $\ell_{\text{LP}} \in [0, 1]$: mix contribution of the low-pass output. At unity, only frequencies below the cutoff pass through; the signal becomes progressively darker.
- **mid** $\ell_{\text{BP}} \in [0, 1]$: mix contribution of the band-pass output. At unity, only the frequency band centred on the cutoff is retained, isolating a tonal or resonant region.
- **high** $\ell_{\text{HP}} \in [0, 1]$: mix contribution of the high-pass output. At unity, only frequencies above the cutoff pass through; the signal becomes progressively brighter and thinner.

- **mix** $m \in [0, 1]$: wet/dry blend applied after the three-mode sum (default $m^{\circ} = 1$, fully filtered).

Combining all three at non-zero weights produces hybrid timbres: equal **low** and **high** with zero **mid** approximates a notch filter, while a dominant **mid** isolates a resonant band without removing the surrounding spectrum entirely. The cutoff frequency is derived from the incoming control signal via a clamp operation,

$$v_k = \text{clamp}(s_d, \ell, h),$$

where s_d is the mapped winch metric and ℓ, h are the operator-configured lower and upper bounds. This constrains the cutoff to a musically useful range without exposing raw filter parameters at the performance interface.

8.4.6 Quadraphonic Panning

The quadPan module distributes the stereo signal across a two-dimensional panning plane using two axis parameters, each derived from a spatial tension relationship between opposing canvas regions:

$$\ell_{\text{LR}} = |\bar{p}_{\text{left}} - \bar{p}_{\text{right}}|, \qquad \ell_{\text{FB}} = |\bar{p}_{\text{front}} - \bar{p}_{\text{back}}|,$$

where $\bar{p}_r$ is the soft-weighted position average of spatial region r as defined in subsection 8.2. Both values range over $[0, 1]$: zero indicates that opposing sides of the fabric are at equal height (symmetric / centred), while values approaching unity indicate maximal height disparity (one side pulled up, the other down). The default is $\ell^{\circ} = 0.5$ (neutral centre). This mapping directly encodes the physical tilt and lean of the fabric into spatial audio position.

8.4.7 Granular Synthesis

The granular module fragments and reassembles the incoming audio using grain-based resynthesis, implemented with `munger~` from the HISSTools package. The module operates in both the standalone Max patch and the Max for Live device. Six parameters control its character:

- **mix** $m \in [0, 1]$: wet/dry blend (default $m^{\circ} = 0$, fully dry; granular is off by default).
- **grainsize** $g \in [0, 1]$: grain duration. Short grains ($g \ll 0.5$) produce pitched artefacts and timbral smearing; longer grains ($g \to 1$) approach the original signal.
- **density** $\rho \in [0, 1]$: number of simultaneously active overlapping grains. Low density produces sparse, stuttering textures; high density approaches continuous resynthesis.
- **spread** $\sigma \in [0, 1]$: stereo scatter of individual grains across the field.
- **jitter** $j \in [0, 1]$: randomness of grain playhead position. Zero jitter produces pure loop-based resynthesis; high jitter samples randomly from the buffer, producing a smeared, cloud-like texture.
- **pitch** $\pi \in [0, 1]$: transposition applied within each grain, centre-encoded at 0.5 (unity).

Mapping winch speed to density and jitter, for instance, produces a texture that grows increasingly fragmented and scattered as the fabric accelerates - a direct sonic analog of mechanical chaos.

8.4.8 Bitcrusher

The bitcrusher introduces lo-fi digital degradation through two independent processes:

- **depth** $d \in [0, 1]$: sample-rate reduction. At $d = 0$ the signal is processed at full rate; as d increases, samples are progressively held longer, introducing the characteristic staircase distortion of downsampling without anti-aliasing.
- **bits** $b \in [0, 1]$: bit-depth reduction. At $b = 0$ the full floating-point precision is retained; increasing b reduces the number of amplitude quantization levels, producing coarse quantization noise and eventual hard clipping. Note that this direction is the inverse of the conventional label convention, where a higher "bit depth" implies higher fidelity. Evaluation sessions did not flag the naming as counterintuitive, and the parameter direction was retained.

The combined operation can be written as:

$$y(t) = Q_b\big(\mathrm{hold}_d(x(t))\big),$$

where hold_d is a sample-and-hold with a hold length proportional to d, and Q_b is a mid-tread quantizer with $2^{B(b)}$ levels, $B(b)$ being a monotonically decreasing function of b. Default values ($d^\circ = 0.45$, $b^\circ = 0.35$) place the module in a mild degradation regime that adds texture without obliterating the signal.

8.4.9 Tremolo

The tremolo module applies sinusoidal amplitude modulation using a dedicated low-frequency oscillator (LFO). Three parameters govern its behaviour:

rate $r \in [0, 1]$: LFO frequency, linearly mapped as $f_{\mathrm{LFO}} = 0.01 + 19.99\,r$ (Hz), covering $0.01\,\mathrm{Hz}$ (one cycle per $100\,\mathrm{s}$, barely perceptible) to $20\,\mathrm{Hz}$ (at the boundary of audio-rate amplitude modulation).

depth $d \in [0, 1]$: modulation depth, controlling the swing of the amplitude envelope.

mix $m \in [0, 1]$: wet/dry blend.

The LFO output $y_{\mathrm{osc}}(t) \in [-1, 1]$ (sinusoidal) is converted to a unipolar envelope:

$$y_{\mathrm{lfo}}(t) = \frac{y_{\mathrm{osc}}(t)}{2} + \frac{1}{2} \;\in\; [0,\, 1].$$

The amplitude envelope with depth modulation is:

$$A(t) = y_{\mathrm{lfo}}(t) \cdot d + (1 - d) \;\in\; [1 - d,\, 1].$$

When $d = 0$, $A = 1$ (no modulation; signal unchanged). When $d = 1$, A oscillates between 0 and 1 (full amplitude modulation). The final output mixes the tremolo-modulated wet path with a dry bypass:

$$\begin{aligned} y_{\mathrm{out}}(t) &= A(t) \cdot m \cdot x(t) + (1 - m) \cdot x(t) \\ &= x(t)\big[A(t) \cdot m + 1 - m\big]. \end{aligned}$$

Both channels share a single LFO, preserving mono phase coherence of the modulation while the audio channels are processed independently.

8.4.10 Ring Modulator

The ring modulator multiplies the input signal by a sinusoidal carrier oscillator, producing sum and difference sidebands. If the input contains a component at frequency f_s, the ring-modulated output contains components at $f_s + f_c$ and $|f_s - f_c|$ but no energy at f_s itself (when the carrier is a pure sine). For harmonic input, this destroys harmonic relationships and creates metallic, bell-like, or inharmonic spectra depending on f_c.

Two parameters govern the effect:

frequency $x \in [0, 1]$: carrier frequency on a logarithmic scale,

$$f_c = 20 \cdot 100^x \text{ Hz},$$

mapping $x = 0 \to f_c = 20\,\mathrm{Hz}$ (subaudible beating), $x \approx 0.22 \to f_c \approx 440\,\mathrm{Hz}$ (concert A, strong inharmonic colouration), and $x = 1 \to f_c = 2000\,\mathrm{Hz}$ (high-frequency metallic shimmer). The logarithmic scale is chosen because equal perceptual steps in carrier pitch correspond to equal multiplicative steps in frequency, matching the structure of Western equal temperament.

mix $m \in [0, 1]$: wet/dry blend. At $m = 0$ the signal passes dry; at $m = 1$ only the ring-modulated signal is heard.

Both channels share the single carrier oscillator, so the same sideband pattern applies uniformly in stereo. The full signal path for one channel is:

$$y_{\mathrm{out}}(t) = m \cdot x(t) \cdot \cos(2\pi f_c t) + (1 - m) \cdot x(t).$$

The logarithmic carrier scale makes it practical to map winch speed or group average directly to the frequency parameter: slow, still fabric produces a low carrier near $20\,\mathrm{Hz}$ (gentle beating), while rapid movement drives the carrier into the hundreds-of-Hz range (metallic, pitched colouration).

8.5 Visualisation: Gaussian Heatmap

The operator interface includes a 2D canvas overlay that renders the distribution of winch activity

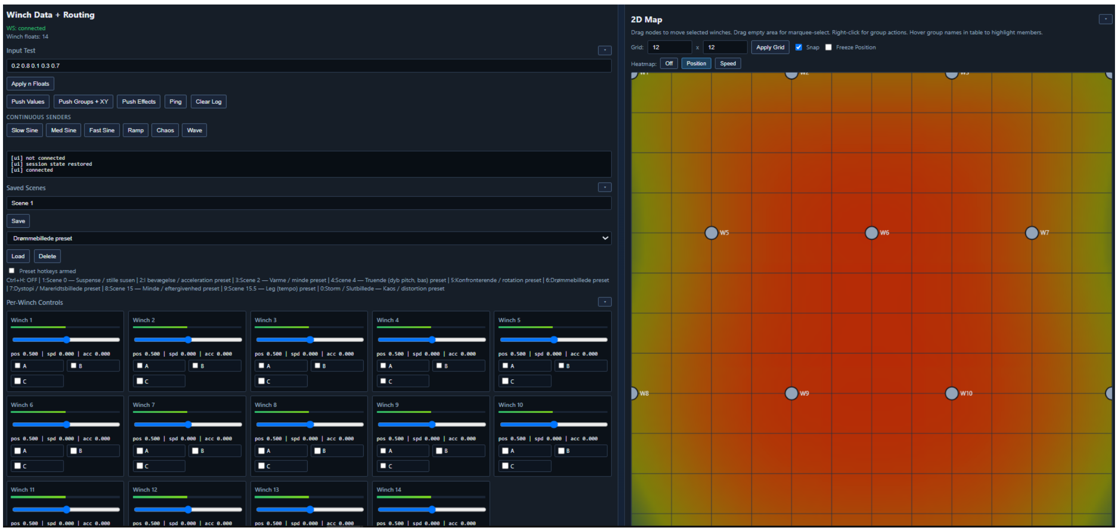

Figure 11: Browser UI: Gaussian heatmap overlay on the winch canvas. Hot regions (yellow–red) indicate high positional activity; cool regions (blue) indicate stillness. The heatmap updates in real time and was identified by evaluators as the system's most effective operational affordance (Q13: 5/5).

as a Gaussian heatmap. At each frame, a scalar field $B(u, v)$ is accumulated over all winches:

$$B(u, v) = \sum_i w_i \cdot \exp\left(-\frac{(u - c_{x,i})^2}{2\sigma_x^2} - \frac{(v - c_{y,i})^2}{2\sigma_y^2}\right)$$

where $(c_{x,i}, c_{y,i})$ are the canvas-space coordinates of winch i, $w_i \in \{p_i, v_i^{\text{abs}}\}$ is the chosen heat metric (position or speed), and $\sigma_x = 0.2 \cdot W_{\text{buf}}$, $\sigma_y = 0.2 \cdot H_{\text{buf}}$ are spread parameters proportional to the buffer dimensions. The field is rendered at $\frac{1}{4}$ canvas resolution for performance, then bilinearly scaled to full size. Values are normalised by the frame maximum before colour mapping:

$$t_i = B(u, v) \,/\, \max_{u,v} B(u, v),$$

and coloured through a three-segment gradient (cool blue → cyan → yellow → hot red) with alpha proportional to t. This provides the operator with an immediate spatial readout of which regions of the fabric are active, moving, or still at any given moment.

9. STANDALONE DEPLOYMENT AND SHOW CONTROL

The Max for Live plugin architecture used during development is well suited to exploratory work but introduces a runtime dependency on Ableton Live that may be undesirable in a permanent theatrical installation. A leaner deployment path is a standalone Max patch that runs the DSP chain without a Live session, receiving OSC directly and routing audio to the venue's PA system without a DAW host. This removes the session-management overhead of Live and reduces the number of running applications the operator must monitor during a performance.

A tighter integration with QLab - the show-control software standard in professional theatrical production - would further reduce operator burden. Rather than requiring manual reconfiguration of mapping presets between scenes, a QLab cue list could trigger preset recall messages over OSC at defined cue boundaries, aligning the sonic layer with the rest of the production's technical cue stack automatically. This would also make it possible to evaluate the mapping framework under more controlled conditions, since the operator's real-time configuration choices would no longer be a confounding variable between winch behaviour and sonic output.

Both paths remain as future development goals. The current implementation prioritises correctness and iterability over deployment minimalism, and the Max for Live workflow is retained for the evaluation phase.

10. EVALUATION

The evaluation addresses two primary dimensions. The first is **metaphorical coherence**: whether the system's sonic output is perceptually legible as the dramaturgical quality it is intended to represent - whether a scene configured for warmth sounds warm, and a scene configured for chaos sounds chaotic. The second is **workflow usability**: whether the browser-based mapping interface supports non-technical operators in configuring and adjusting the system within the constraints of a live theatrical context.

10.1 Evaluation Design

The evaluation was structured in two phases, reflecting the practical uncertainty of access to the winch infrastructure.

Phase 1 - Preliminary evaluation used synthetic motion senders to stand in for real winch data. Each sender generated continuous movement data that activated the effect chain without requiring the winch infrastructure. The motion itself was not choreographed and carried no intended meaning; what was under evaluation was the effect preset design. Each of the twelve scene configurations was presented twice in sequence: first with the effect chain bypassed (the score alone - Version A), then with the effects active and driven by the sender (score plus sonification - Version B). Evaluators were asked to describe what changed between A and B before the intended dramaturgical quality was revealed, then rate how strongly the effect layer shifted the sound towards that quality. The A/B structure isolated the contribution of the sonification from the score itself, and the blind description step tested whether the shift was perceptually legible before any framing was introduced.

Phase 2 - Live evaluation was planned with the full 14-winch array, contingent on access during a rehearsal period. Access was not arranged, and Phase 2 was not conducted.

As Phase 2 was not conducted, the preliminary evaluation constitutes the complete evaluation dataset. The core research question - whether the mapping framework produces perceptually legible sonic expressions - is addressed by Phase 1 alone, given that the motion patterns were designed to activate the same collective gesture parameters the system would compute from live winch data.

10.2 Participants

The preliminary evaluation was conducted by the author. The expert evaluator profile was deliberately broad: production practitioners could assess dramaturgical coherence, and peers with signal-processing literacy could assess whether the sonic character of each pattern is distinguishable and whether the mappings behave as described. Small evaluator numbers are an acknowledged limitation; the evaluation was designed to produce indicative findings rather than statistically significant claims.

10.3 Ethics and Data Handling

All response data is stored locally in the browser's `localStorage` on the evaluator's own machine; nothing is transmitted to a server or third party. Participants export their own responses as JSON or CSV and share these files with the researcher at their own discretion. Audio recordings are made only with explicit verbal consent. Participation is voluntary and participants may withdraw at any point without consequence.

10.4 Evaluation Protocol

Audio playback for the expert session and the Christian and Frederik sessions used a Samsung Radiant360 R5 omnidirectional speaker. Kevin's session was conducted over headphones. The difference in playback format is relevant to the spatial scenes: the R5 produces a stereo field but not a true quadraphonic one, while headphones reproduce the stereo mix directly.

Each evaluation session began with a short contextual primer before any audio demonstration. Participants were shown documentation of *Aerial Waves* - the spatial installation developed by the same interaction designer using the same Wahlberg winch infrastructure - to establish a concrete visual understanding of how the winches move and what the fabric looks like in motion. This ensured that participants could interpret the sonic demonstrations in relation to the physical system they were responding to, without prior exposure to the specific dramaturgical content of *Isnætter*. The primer was found to be an effective contextualisation strategy in the previous iteration of this project, where showing winch documentation before evaluation substantially improved the evaluator's ability to relate sonic behaviour to physical actuation.

For the preliminary phase, the session proceeded as follows:

1. Contextual primer: documentation of *Aerial Waves*, three minutes.
2. For each of the twelve scene configurations in sequence (storyboard scenes sharing a common preset were grouped into a single configuration to keep the session to a manageable length):
   (a) **Version A**: the scene's soundtrack was played with the effect chain bypassed. No sonification was active.
   (b) **Version B**: the same cue was played again with the effect preset active and driven by the synthetic sender.
   (c) The evaluator wrote a free description of what changed between A and B - what the effect layer added or suggested that was absent in the score alone.
   (d) The intended dramaturgical quality (e.g. *Truende*, *Kaos*, *Varme*) was revealed and the evaluator rated on a five-point Likert scale how strongly Version B shifted the sound towards that quality.
3. Nine effect-specific questions assessed the legibility of individual DSP parameters across all demonstrations.
4. Open questions assessed the overall system and suggested improvements.

10.5 Scene-Specific Expression

A central claim of the mapping framework is that different dramaturgical states - warmth versus cold, order versus chaos, stillness versus storm - should produce perceptually distinct sonic expressions. Verifying this claim requires more than confirming that the system produces sound: it requires asking whether the sonification adds a meaningful layer of expression on top of the score, and whether that layer points in the intended direction.

The A/B structure directly tested this. Because evaluators heard the score alone before hearing the score with effects, any shift in perceived quality between the two versions could be attributed to the sonification rather than to the musical material. The evaluation instrument asked evaluators first to describe what changed unprompted, then to rate - after the intended quality was revealed - how strongly Version B moved the sound towards that quality. Twelve scene impressions, derived from the storyboard's own *Effekter* column, covered the full dramaturgical range of the performance. Labels including *Suspense*, *Varme*, *Kaos*, *Dystopi*, and *Slutbillede* were drawn directly from the production's vocabulary.

If evaluators' unprompted A/B descriptions align with the intended quality before the label is revealed, this constitutes evidence that the sonification carries legible metaphorical content. If the Likert ratings are high but the unprompted descriptions do not align, this suggests the shift is present but not specific enough to be identifiable without priming. Both patterns are informative: the evaluation is designed to surface failures as clearly as successes.

11. RESULTS

The preliminary evaluation was completed by four participants across two session formats. The expert session (18 May 2026) was conducted jointly with Thomas Kaufmanas, interaction designer and visual artist with direct involvement in the *Isnætter* production [22], and Tina Tarpgaard, director and choreographer of *Isnætter*. The production's own soundtrack had not been made available at the time of the session; the demonstrations were conducted using *so i will not fall deep into the*

Scene	Intended quality	A/B description (unprompted)	Score
Scene 0	Suspense / stille susen	Metallic, distortion, bevægelse, sharper quality	3/5
Scenes 1, 8, 9, 12, 17	I bevægelse / acceleration	Svinger rundt / pendul, bevægelse, cirkulerende	5/5
Scene 2	Varme / minde	Vakuum, vindforandring, tilbageholdt, volumen	3/5
Scene 4	Truende	Omnious, dystert, farligt, uhyggeligt, klar til ulykke, pitch ned, bevægelse	5/5
Scenes 5, 10b	Konfronterende / rotation	Bevægelse, meget, svingninger, legende	5/5
Scene 7	Drømmebillede (krystal)	Trukket tilbage, distortion	5/5
Scenes 11, 13, 18	Dystopi / pitch høj	-	3/5
Storm / Scenes 18–19	Kaos / storm	Skarp, is, vocalist, clear, dejligt, motion design	5/5

Table 2: Scene-level A/B evaluation results (expert session: Thomas Kaufmanas and Tina Tarpgaard, joint). Descriptions are participants' own words recorded before the intended quality was revealed. Score: 1 = no shift, 5 = strongly shifted towards intended quality.

earth by Ricky Eat Acid as a substitute score, chosen for its thematic and timbral compatibility with the dramaturgical material of *Isnætter*. Three additional independent sessions were conducted with Kevin (academic/SMC background, 17 May 2026), Christian (theatre and production background, 20 May 2026), and Frederik Bjørn (sound design background, 20 May 2026), each using the standard blind A/B protocol with production stems from *Isnætter* spliced together as they would appear in performance. All sessions were audio recorded with consent. As Phase 2 was not conducted, the preliminary evaluation constitutes the complete dataset.

11.1 Sonification Contribution

Table 2 summarises the A/B descriptions and Likert ratings for the eight completed scenes. The *A/B description* column reports participants' unprompted characterisation of what changed between the dry and processed versions, recorded before the intended quality was named; *Score* is the five-point Likert rating of how strongly Version B shifted the sound towards that quality.

Five of the eight scenes received a top rating of 5/5: the movement/acceleration configuration, the threatening configuration (Scene 4), the confrontational/rotation configuration, the crystalline dream configuration (Scene 7), and the storm. Three scenes received 3/5: Scene 0 (Suspense), Scene 2 (Varme/minde), and the dystopian high-pitch configuration.

For the five top-scoring scenes, alignment between the A/B descriptions and the intended quality was strong before the label was revealed. The movement configuration produced descriptions of pendular swinging and circling motion (*svinger rundt / pendul, bevægelse, cirkulerende*), directly matching the intended quality. The threatening configuration prompted descriptions of danger, foreboding, and readiness for catastrophe (*omnious, dystert, farligt, klar til ulykke*). The storm configuration was described as sharp and icy with a quality of deliberate motion design, eliciting the post-reveal note *icy paradise*.

Where scores were moderate, descriptions pointed to concrete revision directions. Scene 0 (Suspense) was characterised as metallic and distorted rather than as the intended airy stillness; participants noted it *should be more airy*. Scene 2 (Varme/minde) was described as a vacuum and a withheld quality (*vakuum, tilbageholdt*), suggesting the effect layer communicated withdrawal rather than warmth. The post-reveal note - *trækker sig tilbage, tilbage til et minde?* (drawing back, back towards a memory?) - suggests partial alignment with the memory dimension of the intended quality but not the warmth. For the dystopian high-pitch configuration, no A/B description was recorded, and participants noted that *pitch down instead of up* would better serve the intended feeling.

11.2 Workflow Usability

Table 3 reports the Likert scores for the sixteen workflow questions grouped by category. Question 14 (mapping assignment controls) was not completed.

Q	Item	Score
Scene Presets		
1	Presets matched storyboard qualities	5/5
2	Switching presets was smooth	2/5
3	Saving and recalling presets was straightforward	4/5
Winch Modulation		
4	Winch movement produced perceptible, meaningful changes	3/5
5	Min/max modulation range is intuitive	5/5
6	Speed and acceleration were responsive as live sources	3/5
Spatial Audio		
7	Winch grid accurately represents stage layout	3/5
8	Quadraphonic panning created meaningful spatial relationship	4/5
Reliability & Feel		
9	UI–Max connection was stable	2/5
10	UI felt fast enough to use as a live instrument	4/5
Interface & Usability		
11	Interface was intuitive without prior instruction	4/5
12	Parameter labels were self-explanatory	4/5
13	Heatmap gave useful collective feedback	5/5
14	Mapping assignment controls were easy to adjust	-
15	Could operate in live rehearsal after short introduction	3/5
Overall		
16	Felt like a purposeful artistic tool	4/5

Table 3: Workflow and usability scores. Five-point Likert scale; 1 = strongly disagree, 5 = strongly agree. Q14 was not completed.

The highest scores were recorded for preset content (Q1: 5/5), the min/max modulation range (Q5: 5/5), and the Gaussian heatmap (Q13: 5/5). The two lowest scores were preset switching smoothness (Q2: 2/5) and connection stability (Q9: 2/5). The Q2 score reflects an evaluation design limitation: the preset system was never explicitly demonstrated during any session. Evaluators discovered it through free exploration but none actively created or loaded presets during the session - the feature remained latent. The score captures first-contact friction rather than a fundamental design failure. The Q9 score corresponds to the session blocker reported in the open response field (*datastream*): the live OSC data connection was not reliably stable during the session, which limited the extent to which live modulation could be demonstrated and directly affected the scores for winch responsiveness (Q4, Q6) and live operability (Q15).

The free exploration response highlighted the heatmap as the system's most effective interface element:

> *"En ting der fungerer virkelig godt er den visuelle fremvisning sammen med lyden, fordi det støtter min forståelse af hvad effekterne gør 'lige nu'. Hvilket ellers kan være svært at høre, da lyd er en langt mere subtil modalitet end visuelle elementer."*
>
> [One thing that works really well is the visual display together with the sound, because it supports my understanding of what the effects are doing 'right now' - which can otherwise be difficult to hear, as sound is a far more subtle modality than visual elements.]

This observation confirms a practical affordance of the combined heatmap-and-audio display that the design had not explicitly anticipated: the visual layer compensates for the perceptual difficulty of monitoring real-time effect modulation by ear alone.

11.3 Preliminary Evaluation

Table 4 reports the cross-evaluator scene scores across all four participants. Scene identifiers correspond to the storyboard numbering used in the mapping framework.

The two strongest configurations were *Truende* (Scene 4, mean 4.75) and *Kaos/Storm* (mean 4.75). Both produced convergent unprompted descriptions across all four participants without priming. Scene 4 was described as dark, foreboding, and dangerous; Kevin wrote *halvt tempo, mørkt, rumlende, mystisk*, and Frederik noted a feeling of "undervandsagtigt hulefornemmelse"

Scene	Intended quality	TT	KE	CH	FR	Mean
Scene 0	Suspense / susen	3	4	3	3	3.25
Scenes 1, 8, 9, 12, 17	Acceleration	5	3	4	2	3.50
Scene 2	Varme / minde	3	4	2	4	3.25
Scene 4	Truende	5	4	5	5	4.75
Scenes 5, 10b	Konfronterende†	5	3	1	3	3.00
Scene 7	Drømmebillede	5	-	4	5	4.67
Scenes 11,13,18	Dystopi	3	5	4	5	4.25
Scene 15.5	Leg / let energi	-	2	1	1	1.33
Scenes 18–19	Kaos / storm	5	4	5	5	4.75

Table 4: Cross-evaluator scene scores. TT = Thomas & Tina (expert session); KE = Kevin; CH = Christian; FR = Frederik Bjørn. Five-point scale; 1 = no shift towards intended quality, 5 = strong shift. - indicates scene not reached. †Konfronterende relies on quad-pan spatial rotation; CH and FR sessions used mono playback, confirmed on record during the FR session; scores reflect infrastructure rather than mapping legibility.

(underwater cave feeling). The storm configuration was described as industrial, crystalline, and inescapable; Frederik associated it with a factory building and a computer game aesthetic, and Tina renamed it spontaneously as *icy paradise*. The *Drømmebillede* and *Dystopi* configurations also scored consistently high (means 4.67 and 4.25). For the dystopia scene, both Kevin and Frederik gave 5/5, with Frederik describing "half-life Alyx vibes" and Kevin noting a "kæmpe forskel, robotisk, static" (huge difference, robotic, static); Thomas and Tina gave 3/5 and proposed that pitch downward rather than upward would better serve the intended claustrophobia.

The *Leg/let energi* configuration (Scene 15.5) scored 1.33 across the three preliminary evaluators - the lowest result of any tested configuration - and was the only scene for which all independent evaluators described the effect as working against the intended quality. Christian noted the processed version felt heavier than the clean version. Frederik identified the underlying mechanism: the musical material itself was not lighthearted, so the tremolo compounded its weight rather than inverting it. This is a music-dependency failure rather than a mapping logic failure. Tremolo scored 4.67 as a standalone effect across the same evaluators (Table 5), confirming its expressive range. The scene preset requires either a different source cue or an effect strategy that reduces perceptual weight - a high-frequency filter sweep or upward pitch shimmer - rather than adding rhythmic agitation to already-heavy material.

Three scenes - Suspense, Varme/Minde, and Konfronterende/Rotation - scored in the 2.33–3.25 range, reflecting setting and infrastructure constraints rather than mapping failures. The Konfronterende preset relied primarily on quad-pan spatial rotation. Christian and Frederik listened on the Samsung R5 (mono point source) and scored it 1 and 3 respectively; Kevin listened on headphones and scored it 3 - suggesting the spatial effect was partially audible in stereo but not legible enough to identify without priming. The Suspense and Warmth presets were audible but operating at the threshold of reliable identification: too present to be invisible, too subtle to be unambiguous.

11.3.1 Effect Legibility

Table 5 reports the nine standalone effect ratings from the three preliminary evaluators, completed after free exploration of the full interface. Quad-pan and granular synthesis are excluded: quad-pan was non-functional in mono; granular synthesis was not implemented in the Max for Live port.

Effect	KE	CH	FR	Mean
Delay	5	4	5	4.67
Tremolo	5	4	5	4.67
Reverb	4	4	5	4.33
Pitch shift	4	4	5	4.33
Filter	4	4	5	4.33
Ring modulation	5	5	3	4.33
Bitcrush	4	5	3	4.00
Granular	*not implemented*			-
Quad-pan	*mono - non-functional*			-

Table 5: Standalone effect legibility scores ($N = 3$ preliminary evaluators). Five-point scale: 1 = did not read as expressive, 5 = immediately legible. Effects with implementation or infrastructure issues excluded from analysis.

Delay and tremolo achieved the highest mean scores (4.67). Reverb, pitch, filter, and ring modulation all scored 4.33. Bitcrush showed mild divergence (Christian: 5, Frederik: 3); Frederik observed that the current preset produces a metallic texture that reads as a new sound layer rather than processing of the source signal. Ring modulation was unfamiliar to Christian but was judged to make sense with movement data. The Thomas and Tina session did not include individual effect ratings, but Thomas's verbal summary of the session aligned with the quantitative pattern: ”det som jeg reagerede mest på var helt klart der, hvor jeg fornemmede en klar pitch, eller når der var distortion, hvor jeg kunne mærke lyden ligesom blive forbrændt” - pitch and distortion were the most immediately legible effect types.

11.3.2 Workflow and Interface

All three preliminary evaluators completed the nine-item usability questionnaire after hands-on exploration. Preset recall (Q3) scored 5/5 across all three participants, confirming the scene-switching mechanism is well-designed. Five specific interface problems were identified across verbal and written feedback.

Slider proliferation. The per-effect panel exposes multiple controls simultaneously - depth, mix, minimum, maximum, and source - for each active effect. Christian wrote ”too many sliders and options made it confusing”; Thomas said ”du kan gøre det nemmere” during the expert session; Frederik described the system as ”lidt overingenieret” and called for ”noget lidt simpler” for live use. This feedback is consistent with the live operability scores: Q15 (live rehearsal after a short introduction) averaged 4.0 across the three preliminary evaluators, but fell to 3/5 for the system overall in the expert session.

Min/max clamping controls. The interface exposes normalisation and clamping controls for every effect source. Both Kevin and Frederik independently identified these as the most cognitively demanding feature of the interface. Kevin observed that ”det er meget sjældent, at man skal have en varierende min maxværdi” (it is very rare to need a varying min/max value) and described the controls as surplus information that creates overload without adding expressive value. Frederik was unable to distinguish the clamping range from the normalisation floor during hands-on exploration. Neither evaluator adjusted the values meaningfully during their session; both defaulted to preset bounds. The controls' function is not communicated by the layout, and their presence competes with the parameters practitioners are more likely to adjust - depth and mix - without proportionate expressive benefit.

Technical naming and effect vocabulary. Effect and parameter labels use DSP terminology (bitcrush, ring modulation, dry/wet) rather than experiential or dramaturgical terms. Frederik suggested that the interface should instead offer experiential group names - terms like ”overdrive” or ”chaos” that each represent a bundle of effects rather than a single parameter - ”nærmere som muligheder” (more as possibilities). He extended this into a structural suggestion: a set of premade effect chains identified by experiential labels - warm, distorted, ambient - that practitioners could select as a starting point before adjusting details. His formulation was ”chains of effects, premade” followed by ”altså igen det der lidt preset” (again that preset idea). This points toward a two-tier interaction model: a curated vocabulary of named sonic states at the top level, with raw DSP controls accessible underneath for those who want them. The current interface offers only the latter.

Normalised-only slider values. All effect parameters are displayed and adjusted as normalised values in the range $[0, 1]$, with no conversion to meaningful units at the interface level. Frederik noted that a pitch shift slider, for instance, gives no indication of the interval it represents in semitones; a reverb time slider communicates nothing about actual decay time in seconds. The absence of unit labels makes it difficult to reason about parameter settings in musical or acoustical terms, and increases the cognitive load of distinguishing between effects that are perceptually similar at a given normalised value.

Heatmap concealed during effect assignment. The heatmap and the effect assignment panel occupy separate views. Frederik noted: "det ville være stort, hvis man kunne se heatmappet, mens man applier effekter" (it would be great to see the heatmap while applying effects). Thomas Kaufmanas confirmed the heatmap as the system's most effective operational feedback element (Q13: 5/5); concealing it during the key configuration step reduces the value of its real-time signal.

11.4 Summary

Across four evaluators in two session formats, the scene-level scores (Table 4) span a wide range. At the high end: Truende and Storm both averaged 4.75/5; Drømmebillede averaged 4.67; Dystopi averaged 4.25. In the middle range: the movement and acceleration scenes averaged 3.50; Suspense (Scene 0) and Varme/Minde (Scene 2) both averaged 3.25; Konfronterende/Rotation averaged 3.00. The leg/tremolo scene (Scene 15.5) averaged 1.33 - the lowest result. The mid-range scores reflect a combination of playback infrastructure constraints (spatial effects inaudible without stereo) and presets operating at the edge of reliable identification. The leg/tremolo result reflects a musical-context mismatch rather than a mapping logic failure: the same effect scored 4.67/5 in isolation.

Effect legibility scores (Table 5) placed delay and tremolo highest (4.67/5 each), followed by pitch shifting (4.33). The workflow evaluation identified the heatmap as the strongest single affordance (Q13: 5/5) and surfaced five specific interface problems documented in subsubsection 11.3.2: slider proliferation, unclear min/max clamping behaviour, technical parameter naming, normalised-only slider values, and the heatmap being hidden during effect assignment. The system was confirmed as a purposeful artistic tool by the expert evaluators (Q16: 4/5).

12. DISCUSSION

The evaluation findings, together with the design and implementation process, surface three recurring themes that bear on the research question: the validity of the dramaturgical grounding strategy, the operability challenges introduced by collective parameters, and the scope limitations imposed by the single-input architecture.

12.1 Metaphorical Legibility and Its Limits

The evaluation provides partial support for the central claim of the mapping framework. Five of the eight tested scenes produced unprompted A/B descriptions that aligned with the intended dramaturgical quality before the label was revealed. The movement configuration was described as pendular and circling; the threatening configuration as dangerous and foreboding; the storm as sharp and icy. These descriptions did not require priming - they emerged from listening alone, suggesting that the embodied metaphors grounding the mappings (SPEED IS TENSION, UP IS HIGH, DISINTEGRATION IS FRAGMENTATION) are perceptually operative in the sonic output. This is consistent with [23], who found in a controlled study that metaphor-grounded sonification of machine telemetry was more intuitively legible to participants than conventional data-driven mappings, even without prior instruction.

The three independent preliminary sessions showed a more varied pattern. The strongest configurations from the expert session - Truende, Drømmebillede, and Storm - held up consistently across all four evaluators. The movement and acceleration scenes, which scored 5/5 in the expert session, averaged 3.50 across the three independent evaluators; the Konfronterende configuration dropped from 5/5 to a mean of 3.00. Both cases are partly explained by playback conditions - neither headphones nor the Samsung R5 reproduced the spatial and quad-pan elements the presets relied on - but the inconsistency across evaluators for the movement scenes also suggests the mapping is less immediately legible without the expert session's stereo environment and shared context. The leg/tremolo scene (Scene 15.5) was the clearest failure in both session formats. The underlying musical material for that scene was heavy and dense; tremolo at a respiratory rate compounded that weight rather than introducing the intended RHYTHM IS LIFE lightness. The effect did not fit this particular musical context - the tremolo effect scored 4.67/5 for embodied legibility when evaluated in isolation (Table 5), suggesting the mapping works but requires lighter source material to land.

The three moderate-scoring scenes in the expert session reveal the limits of the approach. The suspense configuration (Scene 0) was heard as metallic and distorted rather than as the intended airy stillness - a failure of directionality: the mapping activated the correct effect type (ring modulation) but at a density that overshot the target quality. The warmth/memory configuration (Scene 2) communicated withdrawal more than warmth; participants' post-reveal note - *trækker sig tilbage, tilbage til et minde?* - suggests the memory dimension was legible but the warmth dimension was not. This points to a tension within the WARMTH IS INTIMACY metaphor as implemented: the reverb-based spatial withdrawal that encodes closeness at low depths reads as absence rather than presence at the tested settings. The dystopian high-pitch configuration was the most contested: participants judged that pitch downward rather than upward better serves the intended nightmarish quality. For this specific dramaturgical context - violence, claustrophobia, inescapability - embodied gravity (DOWN IS HEAVY, DOWN IS OPPRESSIVE) appears to be a stronger cue than the UP IS HIGH schema the mapping was built on. This points to a general design consideration: where two competing embodied schemas both plausibly apply to a target quality, testing which dominates in context is preferable to committing to either one in advance.

A further form of corroboration emerged after the evaluation was complete. Attending the premiere of *Isnætter* revealed that the production soundtrack independently deploys many of the same sonic gestures - granular fragmentation, delay and echo, spatial processing, pitch transposition, reverberant depth - in the scenes and transitions where the dramaturgical analysis had identified them as appropriate. The score and the system were developed without coordination; their convergence on similar sonic strategies for the same material provides external evidence that the metaphorical framework is tracking something real about the dramatic content, independently of the evaluation results.

12.2 Dramaturgical Grounding as a Design Constraint

A recurring finding across the design process was that grounding mapping decisions in the dramaturgical material - rather than in technical convenience or aesthetic preference - produced choices that were easier to justify, easier to explain to production staff, and more resistant to arbitrary revision. When a mapping can be traced back to a specific annotation in the storyboard and supported by an embodied metaphor from the literature, it occupies a stable position in the design: it can be adjusted in degree but not discarded without losing something that the dramaturgical analysis identified as meaningful. The evaluation confirmed this: the five high-scoring mappings all have clear storyboard anchors and strong embodied metaphor support. The three mappings that failed or underperformed in evaluation are precisely those where the metaphorical grounding was more contested or where a single metaphor was given priority over a competing one.

The collective parameter approach introduced a related constraint. Because the primary sonic parameters - landscape index, kinematic entropy, directional consensus - are derived from the relational state of the full winch array, the sound responds to things that no individual winch controls. This means the operator cannot easily produce a specific sonic result by manipulating a single input; the system must be understood as a whole. This is both the approach's greatest strength - it produces genuinely emergent behaviour - and its primary operability challenge. Operators accustomed to direct control relationships require more time to develop an intuition for how collective configurations translate into audible output. The evaluation score for live operability (Q15: 3/5) reflects this: participants could imagine operating the system after a short introduction, but not without one.

12.3 The Heatmap as an Unexpected Affordance

The highest single workflow score was for the Gaussian heatmap (Q13: 5/5), and participants explicitly identified it as the system's most effective interface element. The observation - that the visual display compensates for the difficulty of monitoring real-time effect modulation by ear alone - points to a design principle that extends beyond this project: in systems where the mapping layer is complex and the sonic output is subtle, a real-time visualisation of the control signal may be as important for operator confidence as it is for debugging. The heatmap was not designed as a primary operational feature; it was included as a diagnostic aid. The evaluation suggests it should be treated as a core affordance in any future iteration.

12.4 Scope and Limitations

The evaluation was conducted across two session formats: an expert session with Thomas and Tina (18 May 2026) and three independent preliminary sessions with Kevin, Christian, and Frederik - giving N=4 evaluators in total, spanning academic, theatre production, and sound design backgrounds. The expert session used a shared response sheet, meaning it represents one joint assessment rather than two independent measurements. Kevin's session used headphones; Christian's and Frederik's sessions used the Samsung Radiant360 R5 omnidirectional speaker, which collapses the stereo field to a single point source. Both playback conditions rendered quad-panning inaudible; this explains the low scores for the spatial-dependent scenes in Table 4 and should be read as an infrastructure constraint rather than a mapping failure. The connection instability in the expert session (Q9: 2/5) similarly reflects session-specific conditions unrelated to the mapping framework.

The expert session was additionally conducted without the production soundtrack, since the *Isnætter* material had not been made available at that point; the legibility findings from that session reflect a substitute score rather than the actual performance context. The three independent sessions used production stems and are not subject to this caveat.

The mapping framework also addresses the winch system in isolation. *Isnætter* deploys lighting, choreography, and the performers' voices and bodies as simultaneous expressive channels, all carrying the same dramaturgical content. In the current implementation, these elements are treated as context rather than input: the sound responds to the winches, and the operator is responsible for ensuring that the sonic result coheres with what is happening elsewhere on stage. This places a significant interpretive burden on the operator and introduces a potential for misalignment between the sonic layer and the live performance that no amount of mapping refinement can fully resolve.

A further physical constraint is that the system models the winch array as a set of attachment point heights and derives all collective metrics from these positions. Fabric draping - the actual surface geometry between attachment points, shaped by textile tension, material properties, and adjacent winch configuration - is not captured by positional data alone. Extended time with the physical apparatus and fabric would more clearly reveal which winch configurations produce perceptually meaningful shapes as mountain peaks, valleys, or travelling waves, and would allow the topological and thermal metaphors to be grounded in observed textile behaviour rather than inferred from attachment point geometry.

13. FUTURE WORK

Several directions emerge directly from the limitations identified above and from the scope constraints of the current implementation.

The most immediate extension of this project is to expand the set of theatrical data sources feeding the mapping system. *Isnætter* deploys lighting, choreography, and moving scenography as simultaneous expressive channels, each carrying the same dramaturgical content through a different medium. The current system listens only to the winches. A natural next step is to incorporate data from the lighting console - intensity levels, colour temperature, cue transitions - as additional mapping sources, allowing the sound to respond to the lighting state directly rather than relying on the operator to maintain coherence between the two layers manually. Similarly, floor-based vibration sensors or camera-based performer tracking could make the choreographic activity of the performers available as a data stream, opening the possibility of mappings that respond to the density, speed, or spatial distribution of bodies on stage rather than fabric in the air.

More broadly, the framework developed here - deriving collective parameters from distributed kinetic systems and grounding mapping choices in embodied metaphor theory - is not specific to winch systems or textile scenography. Any theatrical context that involves multiple independently controlled moving elements generates the relational data that the collective parameter approach requires. Motorised lighting rigs, automated stage platforms, cable-driven scenic elements, and even multi-performer motion capture systems all produce the kind of distributed kinematic streams from which meaningful emergent parameters can be extracted. Applying the framework to these contexts would test both its generalisability and the limits of the dramaturgical grounding approach when the source material is less densely annotated than the *Isnætter* storyboard.

On the interface side, the evaluation surfaced a consistent demand for experiential effect labels - terms like "warm," "stretch," or "chaos" that name a bundle of parameters rather than a single DSP operation. Implementing a two-tier model, with curated named states at the top level and raw controls accessible underneath, would directly address the slider proliferation and naming problems identified in subsubsection 11.3.2 without reducing the system's expressive range.

On the technical side, a closer integration with QLab or a dedicated show control system would reduce the operational burden on the operator during performance, allowing scene-specific configurations to be recalled automatically at cue boundaries rather than adjusted by hand. This would also make it possible to evaluate the system's metaphorical coherence more rigorously, since the operator's configuration choices would no longer be a confounding variable in the relationship between winch behaviour and sonic output. A deeper integration with the Wahlberg control infrastructure itself - receiving machine state directly from the show control network rather than via a forwarded OSC stream - would further reduce latency and remove the dependency on a separate bridging machine during performance. Expanding the evaluation to multiple operators across rehearsal cycles would also strengthen the evidential basis; Reimer et al. [24] provide a transferable framework for longitudinal DMI evaluation that a follow-up study could adopt directly.

The most direct and unresolved future direction is the live deployment of the sonification system in an actual performance of *Isnætter*. The present project delivered a fully implemented, evaluated, and documented system, but the premiere of the production proceeded without its integration. Running the system in the performance space - with the full 14-winch array active, the production soundtrack playing, and an audience present - would be the first opportunity to test whether the mapping framework holds under real conditions, and whether the metaphorical congruence achieved in design is perceptible to listeners who have not been briefed on the system's logic. This is the remaining open question the current evaluation cannot answer, and the one most worth pursuing.

14. CONCLUSION

This project set out to answer whether convincing sonic metaphors can be derived from the data produced by mechanical theatrical set pieces, and whether those metaphors can be made perceptually legible without prior instruction. The answer depends on the mapping strategy. A direct, one-to-one approach - mapping individual winch positions to individual audio parameters - produces output that moves with the machinery but communicates nothing about the world the performance is constructing. The present project pursued a different path: deriving collective emergent metrics from the relational behaviour of the full winch array, and grounding each mapping choice in the dramaturgical vocabulary of *Isnætter* and in the embodied conceptual metaphors that support cross-sensory meaning-making.

The result is SOMATRIC - a system that responds not to individual machines but to the spatial and kinematic character of the array as a whole. Kinematic entropy encodes order and chaos. The landscape index encodes altitude and temperature. Directional consensus encodes wave motion and spatial travel. Each of these parameters is meaningless for a single winch; each emerges only from the relational structure of the full set. The claim of the project is that this class of collective parameter, grounded in dramaturgical analysis and embodied metaphor theory, produces mappings that are both technically robust and artistically meaningful.

This claim is partially supported by the evaluation. Five of the eight tested scenes produced unprompted descriptions that aligned with their intended dramaturgical quality before the label was revealed, confirming that the embodied metaphors operative in the mappings are perceptually legible without instruction. Three scenes produced moderate scores and concrete revision directions: the suspense configuration requires a lighter touch, the warmth configuration communicates withdrawal rather than intimacy at its current settings, and the dystopian high-pitch configuration would be more convincing with a downward pitch shift. These are not failures of the approach - they are refinement targets that the evaluation surfaced precisely as designed. The present report documents the design rationale, implementation, and evaluation findings in sufficient detail that each of these revisions can be pursued in a subsequent iteration.

The most significant limitation is the one the production context imposed rather than one the design introduced: the system was never tested with the full 14-winch array in the performance space, and the premiere proceeded without its integration. This constrains what can be claimed. The evaluation was conducted with synthetic motion senders rather than real winch data, and with two participants in a single joint session rather than independent evaluators. These constraints do not invalidate the findings - the embodied metaphor framework makes predictions that are testable independently of the specific data source - but they do mean that the evaluation addresses the mapping logic rather than the full deployed system. Future work, with access to the winch infrastructure under live conditions, would test whether the metaphorical legibility demonstrated in the evaluation survives the noise, unpredictability, and dramaturgical complexity of an actual performance.

Attending the premiere also revealed that the winch infrastructure was used only in the closing scenes of the performance, following a late creative decision not communicated during the design phase. The production score was found to deploy many of the same sonic gestures the system produces - granular fragmentation, delay and echo, spatial movement, reverberant depth - in precisely the scenes where the dramaturgical analysis had motivated those choices; the score and the system were developed without coordination, and that convergence is the closest thing the project has to independent validation of its central claim.

15. DECLARATION OF AI USE

Artificial intelligence tools were used in the following capacities during this project: The browser-based operator interface described in this report was developed by AI assistance (Claude, Anthropic), which created the implementation of the mock-up JavaScript front end, the Node.js bridge layer, and the OSC communication architecture. AI tools were also used to in literature search and source analysis - identifying relevant work in sonification, embodied cognition, and mapping theory, and helping to locate and summarise specific arguments in cited sources. AI assistance was used for proofreading and editing the report text and debugging the code, as well as inspire and help implement certain effects. The HTML evaluation interfaces used in the evaluation sessions were also implemented with AI assistance; the questions, evaluation flow, and free-response sections were designed by the author.

16. REFERENCES

[1] A. Kavan, *ICE*. London: Peter Owen, 1967.

[2] "Sydhavn Teater — ISNÆTTER." [Online]. Available: https://sydhavnteater.dk/event/isn%C3%A6tter

[3] Wahlberg Motion Design, "Aerial waves," Kinetic installation, exhibited at Sydhavnens Festival and Aarhus Festuge, 2025. [Online]. Available: https://wahlberg.dk

[4] ——, "Winches on world tour with Fred again.." Wahlberg Motion Design blog, Mar. 2026. [Online]. Available: https://wahlberg.dk/blogs/music-cases/winches-on-world-tour-with-fred-again

[5] A. Hunt and R. Kirk, "Mapping Strategies for Musical Performance," 2000. [Online]. Available: https://www.semanticscholar.org/paper/Mapping-Strategies-for-Musical-Performance-Hunt-Kirk/24bae2761b3985db7b43b92553b5cf909c0c5f12

[6] A. Hunt and M. M. Wanderley, "Mapping performer parameters to synthesis engines," *Organised Sound*, vol. 7, no. 2, pp. 97–108, 2002.

[7] G. Lakoff and M. Johnson, *Metaphors We Live By*. Chicago: University of Chicago Press, 1980.

[8] S. Roddy and B. Bridges, "Mapping for meaning: The embodied sonification practitioner and the cognitive metaphor theory," *Journal on Multimodal User Interfaces*, vol. 12, no. 1, pp. 7–20, 2018.

[9] B. N. Walker and G. Kramer, "Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making," in *Ecological Psychoacoustics*, J. G. Neuhoff, Ed. San Diego: Academic Press, 2004, pp. 150–175.

[10] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," *PLOS ONE*, vol. 8, no. 12, p. e82491, 2013.

[11] T. Hermann, A. Hunt, and J. G. Neuhoff, Eds., *The Sonification Handbook*. Berlin: Logos Verlag Berlin GmbH, 2011. [Online]. Available: https://sonification.de/handbook/

[12] D. Smalley, "Spectromorphology: Explaining sound-shapes," *Organised Sound*, vol. 2, no. 2, pp. 107–126, 1997.

[13] M. A. J. Baalman, C. Salter, and D. Moody-Grigsby, "Schwelle: Sensor augmented, adaptive sound design for live theatrical performance," in *Proceedings of the 7th International Conference on New Interfaces for Musical Expression (NIME 2007)*. New York: ACM, 2007, pp. 178–181.

[14] J. Zimmerman, J. Forlizzi, and S. Evenson, "Research through design as a method for interaction design research in HCI," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2007, pp. 493–502.

[15] A. O. Effenberg and G. Schmitz, "Acceleration and deceleration at constant speed: Systematic modulation of motion perception by kinematic sonification," *Annals of the New York Academy of Sciences*, vol. 1425, no. 1, pp. 52–69, 2018.

[16] L. Talmy, "Force dynamics in language and cognition," *Cognitive Science*, vol. 12, no. 1, pp. 49–100, 1988.

[17] C. Spence, "Cross-modal correspondences: A tutorial review," *Attention, Perception, & Psychophysics*, vol. 73, no. 4, pp. 971–995, 2011.

[18] T. Wishart, *Audible Design: A Plain and Easy Introduction to Practical Sound Composition*. York: Orpheus The Pantomime, 1994.

[19] M. E. Wheeler, S. E. Petersen, and R. L. Buckner, "Memory's echo: Vivid remembering reactivates sensory-specific cortex," *Proceedings of the National Academy of Sciences*, vol. 97, no. 20, pp. 11 125–11 129, 2000.

[20] G. Lakoff and M. Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books, 1999.

[21] F. B. Blauenfeldt, "Designing a dynamic music system for narrative audio storytellers," Master's thesis, IT University of Copenhagen, 2025.

[22] T. Kaufmanas, "Thomas kaufmanas portfolio." [Online]. Available: https://portfolio.kaufmanas.dk

[23] J. Simmons, P. Bremner, T. J. Mitchell, A. Bown, and V. McIntosh, "The ballad of the bots: Sonification using cognitive metaphor to support immersed teleoperation of robot teams," *Frontiers in Virtual Reality*, vol. 5, 2024.

[24] P. J. C. Reimer, A. Gupta, C. Guastavino, and M. M. Wanderley, "User experience with digital musical instruments: a transferable method for longitudinal evaluation," *Journal of New Music Research*, 2025.

All material used can e found on https://github.com/MinerMunch/winch-ui/tree/master