

Under what conditions does organized sound generate pre-linguistic meaning?

Andreas Russo

Abstract

Sound reaches us before we have the words to describe it. This paper asks under what conditions organized sound — sound assembled with the intention of conveying meaning, from music to soundscape — produces that felt response without requiring instructions, cultural knowledge, or conscious interpretation. That kind of meaning, registered in the body before it has been named or understood, is what this paper calls pre-linguistic.

The investigation draws on Schopenhauer's argument that music expresses the inner nature of phenomena rather than merely representing it, on Jung's theory of universal pre-cultural predispositions, and on Russell's model of affect as a framework for mapping felt states without cultural interpretation. The method is practice-based: a live quadraphonic sound performance called Stillness/Motion, treated as primary evidence alongside the theoretical frameworks.

The central finding is a single moment: sound rotating across speakers pulled a room of people from different cultures into collective physical movement without instruction or narrative. When the acoustic properties of the sound were distinct enough, the body responded before the mind could ask what it meant, but when they were not, the effect collapsed into passive listening. The conditions under which organized sound generates pre-linguistic meaning seem to be specific, physical, and independent of cultural learning.

Research website: <https://andreasrusso.com/stillness-motion>

Keywords: pre-linguistic meaning; organized sound; spatial sound; embodied response; affect; archetypes; practice-based research

Introduction

My work as a sound designer and composer for visual media repeatedly confronted me with situations in which sound appeared to affect audiences before conscious interpretation occurred. This observation became the starting point for Stillness/Motion, a participatory live performance in which I mixed sound in real time for a room of people, using their movement as a score and my intuition as methodology.

Schopenhauer (1818/1969) argued that music accesses something prior to cultural representation. Jung (1959/2014a) described archetypes as universal, inborn patterns of behavior. This paper is the attempt to examine whether human predisposition to respond to organized sound operates at that same level.

The investigation proceeds through three sub-questions: how does organized sound reach the body before cultural interpretation occurs? How do specific acoustic properties carry that response? What does a practice-based, participatory sound performance reveal about the conditions under which this happens?

The paper proceeds through a literature review situating these frameworks and their relationship to organized sound, a case study of the performance, and a further research section identifying the questions this study opens but cannot answer.

Literature review

Jung (1959/2014a) describes archetypes as pre-existent “forms in the psyche which seem to be present always and everywhere,” stating that “their effects are felt in our most personal life.” Archetypes themselves are irrepresentable, but they produce archetypal images that are culturally shaped and that point directly to the archetype (Jung, 1960/2014b). An example of archetype would be that of the Great Mother, which is not any particular mother or goddess, but the underlying human tendency that produces Isis, Mary, and Kali as its culturally shaped archetypal images. But archetypes extend beyond human figures. An example would be that of the Underworld, which comes from our awareness and fascination with the unknown, death, transformation, and rebirth. This is manifested across cultures as Hades, Duat, and Hel respectively in the Greek, Egyptian, and Norse traditions.

Archetypes are the content of what Jung calls the collective unconscious, a layer of the psyche present across all human experience, not culturally acquired but inborn (1959/2014a). The fact

that no known culture lacks music highlights a universal human predisposition to organize and respond to sound, which appears to be archetype-like, but the fact that organized sound is representable, transmissible, and performable excludes it from being an archetype. However, it is not an archetypal image either since it's not the culturally shaped product of a predisposition, but the medium through which those products are called into being. It triggers a universal response without determining its cultural form.

There seems to be a parallel with what Schopenhauer wrote over a century earlier, where he claims that music is different from all other arts “for it never expresses the phenomenon, but only the inner nature, the in-itself, of every phenomenon.” Not “this or that particular and definite pleasure, this or that affliction, pain, sorrow, horror, gaiety, merriment, or peace of mind, but joy, pain, sorrow, horror, gaiety, merriment, peace of mind themselves, to a certain extent in the abstract, their essential nature” (Schopenhauer, 1818/1969). Where Jung locates this pre-cultural layer in the archetype, Schopenhauer locates it in music itself. Schopenhauer describes what music does, while Jung describes what might be the layer of the psyche it does it to, without either thinker completing the connection.

Jung himself seems to have acknowledged that music was missing from his work. In 1956, pianist Margaret Tilly gave him a private demonstration of the potential of music therapy; he reportedly said that he didn't listen to much music anymore because it deals “with such deep archetypal material, and those who play don't realize this”. After the session, he said he would recommend music therapy in every analysis from there on (McGuire & Hull, 1977/2020), but he never addressed the matter in his writings before his death five years later. That silence is part of what inspired this paper.

Russell's circumplex model of affect (1980) offers a practical complement to Schopenhauer's philosophical claim. Where Schopenhauer describes what music does, Russell provides a framework for mapping the affective states that sound produces: all felt states can be plotted along two axes, valence (pleasant to unpleasant) and arousal (activated to deactivated).

Schaeffer (1966/2017) proposed reduced listening: attending to sound as a pure acoustic object, stripped of its source and cause. The premise is that the listener can strip sound of its meaning, but my performance raises questions about this premise: as the case study will explain, sounds used in Stillness/Motion carried embodied associations regardless of instruction.

Specific acoustic properties seem to operate within that space through physical rather than cultural mechanisms. Juslin and Västfjäll (2008) identify a brain stem reflex, a physiological response triggered by acoustically salient events, including sudden attacks, high frequencies, and spatial movement, before cortical processing has occurred. Spatial behavior, spectral character, and temporal envelope carry pre-linguistic affect precisely because they operate below the threshold at which cultural learning becomes relevant.

Janet Cardiff's multichannel installation *The Forty Part Motet* (2001) demonstrates how sound produces embodied navigation of space before conceptual interpretation occurs. Forty speakers are arranged around a room, each carrying a single voice from a choral recording. Visitors move physically toward the voices that pull them without instruction or cultural cue: whether driven by curiosity or reflex, the movement precedes interpretation.

Case Study

Question

Under what conditions does organized sound generate pre-linguistic meaning — a felt response that precedes instruction, narrative, or cultural framing?

Hypothesis

If the spatial movement of sound is immediate and physical enough, it will produce embodied responses in participants before cognition has time to interpret them. Drawing on Schopenhauer's claim that music expresses the inner nature of phenomena directly and on Russell's circumplex model of affect (1980), the piece hypothesised that sound moving through space could activate felt meaning in bodies without communicating a concept.

Method

Stillness/Motion took place in two back-to-back performances of about eight minutes each, for a total of fifteen to twenty people going in and out of a room equipped with a quadraphonic speaker system and an LFE subwoofer. The original concept was a participatory installation using algorithmic body tracking, in which each participant would be assigned a unique tone, and the proximity to each other would generate harmonic relationships. After months of development, the technical layer proved to be too cumbersome, and panning was too imprecise for people to identify which tone belonged to who, making the sonic consequences of their

movement inaudible. On the advice of professor Adriana Sá, the piece became a live performance: manual fader riding in real time, with zones marked on the floor corresponding to sound layers. When someone was in a zone, that sound was active, and if multiple people were in it, multiple layers of that sound would come up.

The non-environmental sound material was produced through a collective process: four classmates (Autumn Bochart, Justin Enoch, Shyalina Muthumudalige, Irazema Vera) and I gathered around a microphone making non-tonal consonants, breath, percussive mouth and hand sounds. Those recordings, layered with field recordings of wind, rain, and sea waves, formed the texture of the piece. All sounds were chosen for their potential to seamlessly morph into one another. No effects were applied other than minimal equalization. All material was spread across the quad system, with some only feeding the subwoofer. A Max/MSP patch allowed real-time mixing, placement, and rotation of each sound across the speakers at variable speeds.

Results

The correlation between zone occupation and sound layers was largely illegible. Without a strong initial discovery moment (the unambiguous physical realization of how each person was affecting the soundscape) most participants defaulted to passive listening. The instruction given at the door, that they were free to walk, dance, or stand still, seems to have communicated freedom without communicating consequence.

Two moments of embodied response came near the end of the first performance, when I tried to lead people. Rotating a wind sound through the four speakers in an accelerating motion led one participant to spin in place, and then all four participants to walk in a circle following the direction of the sound. As the sound rotation accelerated so did people, and within seconds everyone in the room was running in a circle, laughing. When I introduced sounds of tapping on a wooden box reminiscent of fast footsteps, one participant began tip-tapping her feet, while still circling around.

As people were still spinning, a technical failure occurred: the vibration from their stomping interrupted the connection between computer and audio interface, cutting the sound completely. The room stopped abruptly, with people looking around with uncertainty, until they started clapping and laughing again.

Discussion

The circling moment is the paper's central finding. True to Schopenhauer's argument, the rotating wind sound carried no pre-established meaning, yet participants followed its motion with their bodies. The same mechanism seemed to take place with the wooden box, when a granular, transient sound with the acoustic profile of footsteps caused physical mirroring: the body responded, whether by conscious imitation or reflex. In both cases, what Juslin and Västfjäll would identify as a brain stem reflex seemed to precede any conscious decision. In the case of the rotating wind specifically, spatial movement appeared to be decisive: it worked where static, undifferentiated layers did not, suggesting that directional sound tracking movement through space is what triggered the physical response. The technical failure may reinforce this: when the sound cut, the collective movement stopped instantaneously, suggesting the physical response was being actively sustained by the sound rather than by social dynamics alone.

We don't know what the spinning moment meant to each person, but what is observable is that the response was collective, physical, and required no cultural knowledge to produce: a room of people from several different countries, with no shared instruction, responded collectively to a rotating sound. This is consistent with what Jung's framework describes — a layer of response that precedes cultural shaping — without confirming which archetype, if any, was activated. If the performance makes anything legible, it is the mechanism of people's reaction, not the meaning.

Russell's framework seems to be partially confirmed: arousal was clear, as the accelerating rotation of the wind seemed to produce an obvious escalation in the room, while valence was not, since the piece never became legible enough as “emotional geography” for the participants to navigate. How specific acoustic properties produce these effects independently of cultural learning remains an open question this paper identifies but cannot resolve.

Conclusion

The performance partially confirmed the hypothesis. When the relationship between sound and physical response was direct enough to bypass interpretation, pre-linguistic meaning was transmitted and collective. When it was not (which was most of the performance) participants retreated into passive listening. The failure is as informative as the success: organized sound

generates pre-linguistic meaning only when the perceptual chain is immediate enough that the body responds before the mind asks what it means.

Further Research

The most immediate gap this study leaves open is the internal experience of the participants. Reports of their own felt experience collected immediately after the performance would be the logical solution.

For a more nuanced control over the performance, it could be worth considering employing multiple performers, a reduced number of participants, and instrumentation with a broader range of timbre, pitch, and dynamics.

The presence of a visible, active performer shaping the sound in real time is a variable this study can't isolate. Whether the results would differ with a genuinely autonomous system where participants understood they alone were producing the sound might be worth testing in a future iteration.

Finally, the relationship between spatial sound design, proprioception, and pre-linguistic affect in narrative film could be a natural next step for this research.

References

Schopenhauer, A. (1969). *The World As Will and Representation* (E. F. J. Payne, Trans.).

(Original work published 1818) Dover Publications, Inc.

Jung, C. G. (2014a). *The Collected Works of C. G. Jung, volume 9 (part 1): Archetypes and the Collective Unconscious*. (Original work published 1959) Princeton University Press.

Jung, C. G. (2014b). *Collected works of C. G. Jung, volume 8: The Structure and Dynamics of the Psyche*. (Original work published 1960) Princeton University Press.

McGuire, W., & Hull, R. F. C. (Eds.). (2020). *C.G. Jung Speaking: Interviews and Encounters*.

(Original work published 1977) Princeton University Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social*

Psychology, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>

Schaeffer, P. (2017). *Treatise on musical objects: an essay across disciplines* (C. North & J. Dack, Trans.). (Original work published 1966) Univ of California Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(6), 559–575.
<https://doi.org/10.1017/s0140525x08006079>

Cardiff, J. (2001). *The Forty Part Motet* [Sound installation]. Various venues.