

The Aesthetics of AI Error Sounds:
A Spectromorphological Analysis of Vicky Clarke's *Aura Machine*

Tianrui Xiong

1. Introduction

Artificial intelligence (AI) systems for electroacoustic music generation typically learn from source sound materials and generate new sounds that retain perceptible bonding to their sources.¹ Due to their internal complexity, however, such systems can generate outputs that deviate from expectations and are often treated as errors requiring correction.² This raises the question of how AI-generated error sounds behave when employed as compositional materials rather than corrected, and what structural functions they may perform within an electroacoustic piece. For consistency, this essay uses the term ‘AI error sounds’ to refer to such outputs.

However, existing literature provides limited analytical discussion of this issue. To address this gap, this essay takes Vicky Clarke’s *Aura Machine*³—an approximately ten-minute stereo fixed-media acousmatic work, premiered at RNCM PRiSM’s *Future Music 3: Unsupervised*—as an analytical case study. Clarke has since developed live and album versions of the project, but this essay focuses on the original fixed-media version. The work is selected because it foregrounds AI error sounds as a core compositional method and provides extensive public documentation.⁴

Clarke is a Manchester-based sound artist working with sound in relation to human–machine systems.⁵ During her NOVARS and PRiSM residencies, Clarke began treating AI error sounds as primary compositional material in *Aura Machine*.⁶

The sounds of the work are composed of Clarke’s self-recorded source sounds and AI-generated sounds produced by an AI system, whose outputs retain perceptible bonding to the source sounds and are characterised by a rough-edged, lo-fi quality. Notably, however, Clarke did not limit herself to the expected AI source-bonded sounds generated by the system. She deliberately fed the system ambiguous sound materials to provoke the generation of error sounds, whose perceptual bonding to the source sounds is only faintly retained.⁷ Such error

¹ Antoine Caillon and Philippe Esling, ‘RAVE: A Variational Autoencoder for Fast and High-Quality Neural Audio Synthesis’, *arXiv preprint*, arXiv:2111.05011 (2021).

² Zengqiang Shang et al., ‘Analysis and Solution to Aliasing Artifacts in Neural Waveform Generation Models’, *Applied Acoustics*, 203 (2023), Article 109183.

³ Vicky Clarke, ‘AURA MACHINE’, *RNCM PRiSM Blog*, <https://www.rncm.ac.uk/research/research-activity/research-centres-rncm/prism/prism-blog/aura-machine/> (accessed 4 December 2025).

⁴ Vicky Clarke, ‘AURA MACHINE’, *AURA MACHINE*, <https://auramachine.blog/> (accessed 7 December 2025).

⁵ Vicky Clarke, ‘About’, *Vicky Clarke*, <https://vickyclarke.org/about/> (accessed 7 December 2025).

⁶ Clarke, ‘AURA MACHINE’ (*RNCM PRiSM Blog*).

⁷ Ibid.

outputs constitute the AI error sounds examined in this essay. Consequently, Clarke organises these sound materials temporally to sonify the system's Input–Training–Output operational workflow.⁸ Figure 1 illustrates this process.

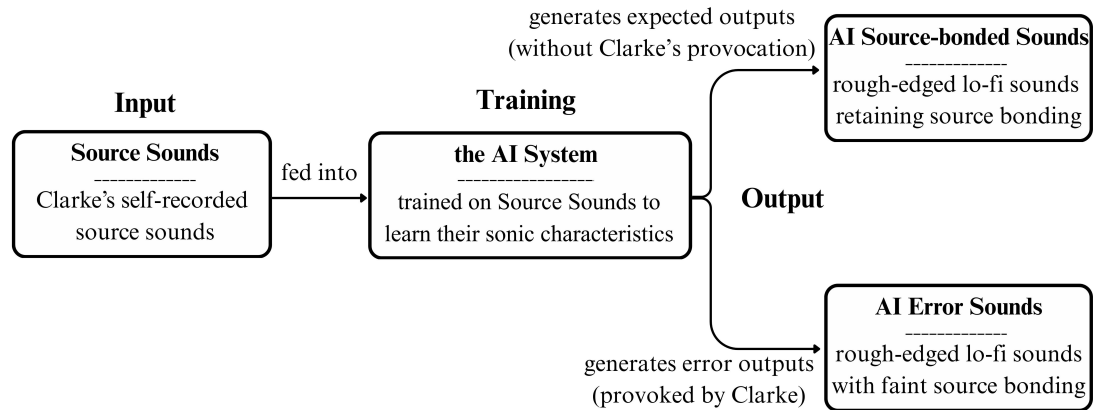


Figure 1. Input–Training–Output workflow of the AI system in *Aura Machine*.

Based on the background above, the sounds in *Aura Machine* can be categorised into three groups: source sounds, AI source-bonded sounds, and AI error sounds. The latter two are lo-fi and rough-edged, perceptually distinguishing them from the source sounds, and are derived from the source sounds while exhibiting different degrees of ‘source bonding’.⁹

Accordingly, this essay uses the framework outlined above to situate AI error sounds within the work's complete sound materials and overall organisation before examining their behaviours and structural functions in detail. The following section outlines the methodology adopted in the analysis.

2. Analytical Methodology

The analysis builds on a temporal listening-based approach and draws on Denis Smalley's spectromorphology¹⁰ as a framework for understanding how sounds behave and function over time.

The temporal listening-based approach is adopted because the overall structure of *Aura Machine* sonifies the AI system's Input–Training–Output process through the temporal organisation of the three sound types: source sounds, AI source-bonded sounds, and AI error sounds. In analysing sound behaviours and structural functions within this temporal flux,

⁸ Clarke, ‘AURA MACHINE’ (*RNCMPRiSM Blog*).

⁹ Denis Smalley, ‘Spectromorphology: explaining sound-shapes’, *Organised Sound*, 2 (1997), 2: 107–126; here, 110.

¹⁰ *Ibid.*, 107–126.

Smalley foregrounds a listening-based understanding of spectromorphological characteristics while deliberately ‘ignoring technology’.¹¹ Within the three sound types, AI source-bonded sounds and AI error sounds—the explicitly technology-based categories—may seem to challenge this position.

In *Aura Machine*, all three sound types are perceptually distinguishable through source bonding, a spectromorphological distinction that remains salient across the work. This suggests that they can be understood in terms of their spectromorphological characteristics rather than solely through technological explanation. This perspective is further supported by András Blazsek’s argument that technologically produced sounds can be treated as analysable sonic objects through auditory perception.¹²

On this source bonding-based foundation, the analysis applies spectromorphological listening to examine sound behaviours and structural functions across different sound types, while reduced listening is used to focus on the intrinsic spectromorphological characteristics of each type.¹³ Specifically, gesture–texture and motion–growth are used to analyse forming principles and temporal development as behavioural processes, while onset–continuant–termination is used to examine how such behaviours produce structural functions within the work.¹⁴

Building on the strategies above, the analytical process is informed by Andrew Lewis’s analytical study of Francis Dhomont’s *Novars*.¹⁵ Following a similar procedure to his analysis of the Golden Section, the analysis traces how a specific compositional element—AI error sounds—behaves and functions within the work’s structure.

The sound materials are first classified and labelled according to source bonding through repeated close listening (Figure 2). These sound types are then mapped onto a temporal structural plan (Figure 3). For each category, descriptions grounded in reduced listening are provided to outline their characteristic spectromorphological features.

Following these analytical steps, a time-referenced graphic score is produced (see Appendix), in which AI error sounds are explicitly marked. Taken together, these reference materials support a detailed, verifiable spectromorphological analysis of the work’s overall structural progression, enabling a perceptually derived structure (Figure 4). This perceptual

¹¹ Smalley, ‘Spectromorphology’, 107–109.

¹² András Blazsek, ‘Machine Listening as Sonification’, *Organised Sound*, 29 (2024), 3: 275–284; here, 275.

¹³ Smalley, ‘Spectromorphology’, 111 (citing Schaeffer).

¹⁴ *Ibid.*, 112–13, 115 (onset–continuant–termination), 113–14 (gesture–texture), 115–17 (motion–growth).

¹⁵ Andrew Lewis, ‘Francis Dhomont’s *Novars*’, *Journal of New Music Research*, 27 (1998), 1–2: 67–83.

structure is subsequently related to the work's form and narrative, with corresponding mappings provided (Figure 5).

Finally, drawing on these analyses and a quantitative comparison of the temporal proportions of each sound type (Figure 6), the analysis concludes with a focused discussion of the behaviours of AI error sounds and the specific structural functions they perform within the work's form and narrative.

3. Analysis of *Aura Machine*

3.1 Sound Classification

To investigate AI error sounds in *Aura Machine*, the first step is to classify the complete sound materials of the work using a set of defined labels, allowing AI error sounds to be distinguished from other sound types. This classification employs repeated close listening and two classification dimensions grounded in source bonding.

The first perceptual distinction, as outlined earlier, concerns degrees of source bonding. On this basis, the work's sound materials can be grouped into three bonding states: source sounds, AI source-bonded sounds, and AI error sounds.

A second dimension is introduced concerning the source domain. Following Clarke's organisation of her source materials, the sounds in the work are grouped into 'Age of Electricity', 'Echoes of Industry', and 'Materiality', corresponding broadly to electrical, industrial, and metal or glass sounds. Taken together, each source domain contains three bonding states, and each bonding state is perceived across all three source domains. To avoid the limitations of a single-dimensional classification, the analysis combines source bonding states (*a–c*) and source domains (*1–3*) to construct a 3×3 classification table of nine sound types (*a1–c3*) in Figure 2. This classification provides a shared analytical foundation for the subsequent analysis. To facilitate identification, all sound-type labels (*a1–c3*) are presented in italics in this essay.

Bonding States Source Domains	<i>a</i> : Source Sounds	<i>b</i> : AI Source-bonded Sounds	<i>c</i> : AI Error Sounds
1. Electrical Sounds	<i>a1</i>	<i>b1</i>	<i>c1</i>
2. Industrial Sounds	<i>a2</i>	<i>b2</i>	<i>c2</i>
3. Metal or Glass Sounds	<i>a3</i>	<i>b3</i>	<i>c3</i>

Figure 2. Sound classification table.

Figure 3 presents a detailed temporal structural plan of all classified sounds across the work. The following descriptions then summarise their principal perceptual characteristics, informed by reduced listening, together with their approximate temporal distributions.

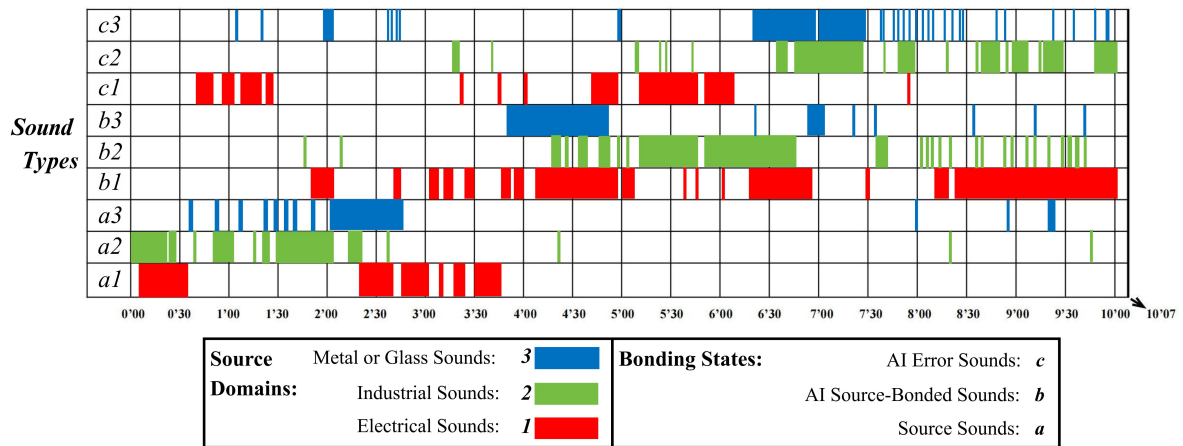


Figure 3. Structural plan of classified sounds.

(It should be noted that some sounds may technically be associated with more than one source or bonding feature. However, the classification is based on the most salient perceptual characteristics through listening, which does not affect the analysis of AI error sounds in this essay.)

Source Sounds (Group a)

Representing the stable physical world, often characterised by clear maintenance and texture-setting behaviours:

a1—Source Electrical Sounds: strongly bonded, texture-carried electrical continuants, often exhibiting sustained oscillatory motion (0'05–0'35; 2'21–4'00).

a2—Source Industrial Sounds: strongly bonded industrial materials establishing iterative, texture-setting planes or entering into flocking behaviour (0'00–0'23; 1'27–2'05; 8'21).

a3—Source Metal or Glass Sounds: strongly bonded impact or fragment gestures that articulate closure, or sustain as flowing texture-carried layers (0'35–1'20; 2'08–2'21; 7'59; 9'22).

AI Source-bonded Sounds (Group b)

AI-generated variations that often function as transitional elements, characterised by continuous flow and specific gestural trajectories:

b1—AI Source-bonded Electrical Sounds: texture-carried electrical layers characterised by continuous flow, or gesture-framed sounds that articulate closure (1'50–2'08; 2'21–4'00; 4'08–5'00; 5'46; 6'18–6'45; 7'28; 8'21–10'00).

b2—AI Source-bonded Industrial Sounds: industrial sounds exhibiting unidirectional descent, divergence, or maintaining streaming texture-settings (2'08; 4'16; 4'24; 5'00–5'46; 6'04–6'45; 8'00–8'21).

b3—AI Source-bonded Metal or Glass Sounds: flowing texture-carried presence or gesture-carried attacks with extended dissipation (3'48–5'00; 7'33–8'00).

AI Error Sounds (Group c)

AI error behaviours characterised by instability, intensified endogeny, and erratic granulated motion:

c1—AI Error Electrical Sounds: fine-grained micro-noise undergoing rapid agglomeration and endogeny, or forming erratic, granulated planes (0'40–0'50; 0'55–1'19; 1'23; 3'17; 3'42; 4'42–5'00; 5'08–5'46; 5'50–6'10).

c2—AI Error Industrial Sounds: texture-carried friction sounds undergoing intensified endogeny to form granulated bodies or entering into convolution (5'08–5'10; 6'35–7'28; 7'48–7'58; 8'35–9'30; 9'47–10'00).

c3—AI Error Metal or Glass Sounds: impacts that condense through agglomeration into thick, blurred forms, or combine to create erratic, iterative turbulence (1'04; 1'19; 1'57; 2'38; 5'00; 6'26–7'28; 7'33; 8'02–8'21; 8'26–9'48).

3.2 Graphic Score

With the analysis outlined above, the analysis therefore constructs a graphic score that represents the behaviour and distribution of all classified sounds over time.

The graphic score (see Appendix) follows Manuella Blackburn's use of spectromorphological visual sound-shapes as 'snap-shots of spectral space' and as a graphical vocabulary for displaying spectromorphological characteristics.¹⁶ Time is mapped on the horizontal axis, and frequency on the vertical axis, while icon shape, shading, and size indicate sound type, textural density, and perceived loudness, respectively.

In line with Blackburn's emphasis on establishing individual visual approaches to spectromorphological representation,¹⁷ boundary outlines are introduced to distinguish the source bonding states of each icon: solid outlines for Group a, dashed outlines for Group b, and cracked outlines for Group c. In addition, all sounds in Group c are explicitly highlighted in the graphic score to facilitate their identification and subsequent analysis.

3.3 Structural Analysis

Combined with the analytical framework and graphic references outlined above, the analysis then adopts a section-by-section temporal listening approach grounded in reduced listening to provide a detailed analysis of the behaviours and structural functions of the classified sound materials. Each section corresponds to a perceptually coherent spectral field, with section boundaries identified when clear changes in the work's overall spectral organisation occur. Figure 4 illustrates this perceptual structure.

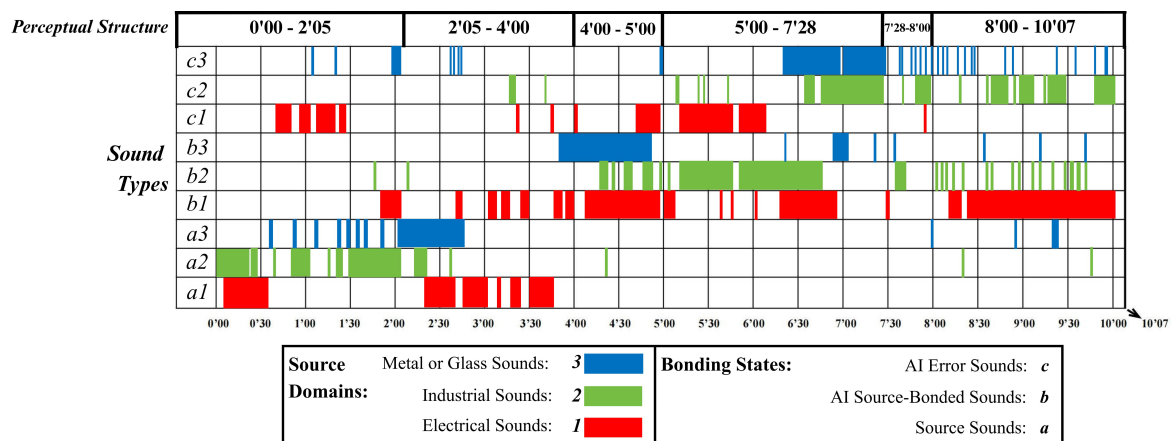


Figure 4. Perceptual structure.

¹⁶ Manuella Blackburn, 'The Visual Sound-Shapes of Spectromorphology: an illustrative guide to composition', *Organised Sound*, 16 (2011), 1: 5–13; here, 5.

¹⁷ Blackburn, 'The Visual Sound-Shapes', 13.

In the analytical description, gesture–texture, motion–growth, and onset–continuant–termination are applied as primary perspectives to trace how the classified sounds behave, transform, and articulate structural functions across the work. In this section, AI error sounds receive a more detailed analytical treatment, in line with the analysis’s focus.

0’00 - 2’05

The work starts with mid-band *a2* industrial pipe-tapping and pipe-friction sounds. These two sounds enter into a flocking behaviour, together establishing an iterative, texture-setting plane. At 0’05, a texture-carried upper-band *a1* electrical hum enters, adding sustained oscillatory motion to the ongoing texture. Later, a group of lower-band gesture-carried *a2* industrial pipe-tapping sounds intermittently enter the texture-carried spectrum, bolstering a stable listening condition. From 0’23, the mid and lower layers move into release, while the *a1* electrical hum undergoes two stepwise, unidirectional spectral ascents, pushing the listening to a peak at 0’35, articulated by a large low-band *a3* metal impact as closure.

This impact persists as a long dissipation trajectory until 0’47. During this period, AI error sounds first appear at 0’40: upper-band *c1* electrical sound enters as fine-grained micro-noise and rapidly undergoes agglomeration. Its internal texture takes on an accelerating, erratic, granulated motion that thickens through endogeny and frames a strongly directional gestural force, generating forward pressure to push temporal progression toward 0’50, where the *a3* metal impact articulates closure once more.

Such a pattern of *c1* micro-noise agglomeration transitioning toward a metal impact is repeated twice between 0’50 and 1’20. In contrast to the first instance, in the latter two progressions, the *c1* micro-noise remains in continuous endogenic transition and agglomeration, condensing into short, lower-band, thick *c3* metal impacts, before the *a3* metal impact sound articulates closure.

At 1’23, following the dissipation of the *a3* metal impact, a gesture-carried *c1* electrical sound occurs as erratic, non-directional motion and passes through the spectrum. Then, from 1’27, a mid-band *a2* industrial friction sound and several periodic *a3* metal impacts continue in maintenance. At 1’50, an upper-band *b1* electrical hum emerges as a texture-carried layer, characterised by continuous flow. From 1’57, the lower-band *c3* metal impacts re-enter and interact with the *a2* friction sound, which has become increasingly thick through abrupt endogenic growth. Together, these sounds form an erratic, iterative turbulence that pushes temporal progression toward 2’05, where a sense of arrival is reached.

2'05 - 4'00

This arrival is emphasised by a short *b2* industrial machine sound that behaves as a rapid unidirectional descent at 2'08, strongly articulating termination and leading into a new spectral field.

After 2'08, with the flow of the upper-band texture-carried *a3* glass fragments, the mid- and low-band *a2* and *b2* sounds function as a transition, leading toward the emergence of an *a1* texture-carried mid-band electrical pad. Between 2'21 and 4'00, this *a1* pad and a *b1* texture-carried lower-band electrical pad engage in a sustained, periodic transition. The process cycles between dilation and contraction of the spectral body, producing a long-timescale oscillatory listening impression.

During this process, AI error sounds function as transitional elements linking the alternation between the erratic, iterative *c3* metal impacts at 2'38, and *c1* spiral, fly-like electrical sounds at 3'17 and 3'42. At 4'00, the *c1* spiral electrical sound articulates the closure of the pad layer. At this point, an upper-band *b3* glass-fragment layer, emerging from 3'48 and maintained as a flowing texture-carried presence, becomes the sole occupant of the spectrum, carrying the sound-field into a new spectral body.

4'00-5'00

From 4'00, the *b3* glass-fragment layer continues in maintenance, sustaining a flowing texture-setting presence. At 4'08, a mid-band texture-carried *b1* electrical hum emerges and joins the *b3* layer, and together they maintain a continuous flow until close to 5'00, before moving into release.

Within this texture-setting field, gesture-carried *b2* industrial machine sounds appear twice at 4'16 and 4'24. These briefly adopt plane-like behaviour, drifting and spreading outward through divergence, before dissolving back into the sustained texture.

At 4'42, the *c1* electrical micro-noise re-enters. Unlike earlier occurrences, its granulated activity proceeds at a slower rate, forming a stable flow rather than undergoing agglomeration. This micro-noise does not form gestural force, and instead, it functions as a texture-carried transitional layer, mediating the upper-band *b3* glass fragments.

At 5'00, both the *c1* micro-noise and the mid-band electrical hum move into release and disappear. A very brief, rapid, periodic, iterative *c3* metal impact follows, articulating closure of the preceding field while simultaneously functioning as an onset into a new spectral field.

5'00 - 7'28

Immediately after this *c3* impact, a low-band texture-carried *b2* plane enters at 5'00 and sustains briefly in maintenance as a fast, dense, continuous layer. At 5'08, this plane is abruptly interrupted by a short *c2* texture-carried industrial pipe friction sound. Unlike earlier pipe-friction sounds, it undergoes intensified endogeny, thickening into a granulated body. It also generates a forward-driving gestural force that cuts off the preceding state and pushes temporal progression forward.

At the same time, an upper-band *c1* electrical micro-noise creates a fast, granulated plane, which is followed by a texture-carried low-band *b2* industrial wind. Together, the two layers form a streaming texture-setting and maintain a continuous flow, within which brief *b2* industrial friction gestures intermittently appear. At 5'46, a *b1* gesture-framed electrical sound articulates closure, bringing the preceding texture-setting to an end. After a short pause, the same setting re-emerges at 5'50 and enters maintenance.

From 6'04, this setting is gradually transitioned into a new texture-setting, formed by the emergence of a mid-band *b2* industrial pipe-tapping sound at 6'04 and an upper-band *b1* electrical hum at 6'18. As this setting develops, it slowly transitions again into another texture-setting, composed of low-band *c3* metal impacts and upper-band *c2* industrial pipe-friction.

The low-band *c3* metal impacts are totally different from earlier appearances: through intensified endogeny and agglomeration, they become highly blurred and granulated, functioning as a ground-layer texture rather than iterative impacts. Above this, the upper-band *c2* industrial pipe friction sustains an erratic, granulated flow.

By 7'00, the low-band impact texture converges and contracts. Each impact becomes increasingly articulated and iterative, forming a forward-driving gestural pressure that pushes the work toward an arrival at 7'28. This arrival is immediately followed by a *b1* gesture-framed electrical sound, which articulates closure and opens a new spectral field.

7'28 - 8'00

At 7'33, a gesture-carried *b3* metal impact with an extended dissipation starts a new spectral field. Following this, *c3* spiral fly-like glass fragments and *c2* industrial pipe-tapping sounds intermittently occur in the field.

At 7'48, the gesture-framed *c2* industrial friction sound reappears, generating a forward-driving force that pushes the work toward 7'56, where another emergence of fly-like glass fragments occurs. This is immediately followed by a *c3* metal impact gesture, which is

tightly coupled with the long dissipation of a *b3* metal impact rather than the dissipation of the *c3* impact itself. Together, this *c3–b3* combination forms a distinctive composite listening impression that closes the previous field at 8'00 and opens a new spectral field.

8'00 - 10'07

From 8'00, a gesture combination recurs multiple times: a mid-band *b2* industrial pipe friction sound immediately followed by a short *c3* metal impact at 8'02, 8'06, 8'13, and 8'21. Unlike the *a3* metal impacts, these *c3* impacts have undergone endogeny, resulting in a thicker spectral body.

At 8'21, following a brief burst of *a2* industrial pipe-tapping, this *b2–c3* combination leads to a low-band *b1* electrical hum. This *b1* layer then establishes a sustained plane that maintains until 10'00.

Between 8'21 and 10'00, isolated *c3* metal impacts at 8'26, 8'52, 9'36, 9'40, 9'44, and 9'48 periodically enter the low band of the spectral field. In the upper band, gesture-framed *c2* industrial pipe frictions recur periodically between 8'35 and 9'30, each briefly emerging and releasing at roughly one-second intervals. Other elements, including *b1* electrical hum sounds, also recur at relatively regular intervals, evoking a feeling of an organised temporal listening field.

At 9'48, the thick lower-band *c2* industrial friction sound reappears, alongside the final *c3* metal impact. The *c2* friction sound enters into convolution with the sustained low-band *b1* electrical hum, creating a mixed textural motion that drives the work's progression toward a final arrival at 10'00, where the overall sound field dissipates.

After approximately seven seconds of silence, the work reaches its definitive termination.

3.4 Form and Narrative

The structural analysis has presented six perceptual sections of the work. Clarke, however, explains the overall progression of *Aura Machine* as a narrative that sonifies the AI system's operational workflow, moving from a stable input, through an unstable training phase, to a stabilised output. How, then, do these six perceptual sections combine to form a large-scale structure that embodies the Input–Training–Output process, and what role do AI error sounds play within this trajectory?

This section maps the work's six perceptually defined sections onto the Input–Training–Output framework, providing a basis for understanding the relationship between the

work's formal organisation and its narrative trajectory (Figure 5). It also establishes a basis for the subsequent discussion of how AI error sounds contribute to the work's structure.

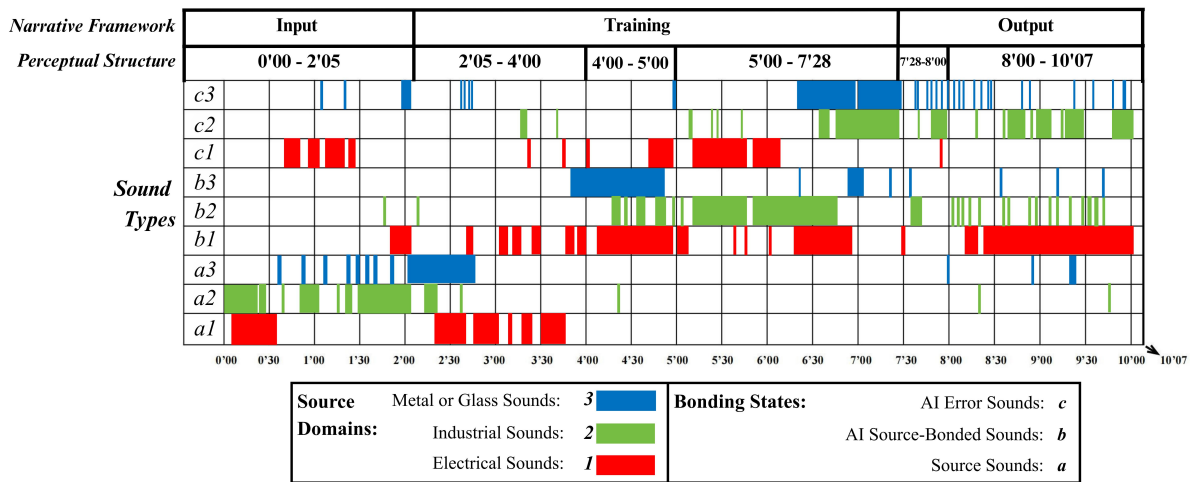


Figure 5. Input–Training–Output narrative mapping.

Input (0'00–2'05)

This phase functions as the exposition of the source sounds (Group *a*). The overall development of this phase begins with a maintained texture-setting field in the mid-to-high spectral bands, formed by *a2* industrial sounds and *a1* electrical hums. This field is followed by recurring low-band *a3* metal impacts, and the phase is driven toward its conclusion by a low-band textural turbulence that pushes the spectral field to an arrival.

This section is defined as the input phase because it is dominated almost exclusively by the source sounds (Group *a*), with maintenance as the prevailing structural function. Narratively, it presents the input stage as a stable sonic representation of a physical source world, in which source sounds are introduced as the primary elements for the AI system.

Training (2'05–7'28)

This phase presents the transmutation of all nine sound types (Groups *a1–c3*). It opens with periodic transitions between the *a1* and *b1* electrical pads, and gradually expands into denser networks of transitions both between and within different sound types. Toward the end of the phase, low-band *c3* metal impacts generate forward-driving pressure that pushes the spectral field toward a *b1* electrical sound, articulating a closure.

Structurally, this phase is characterised by a spectral field in which transition becomes the dominant structural function, as sounds from different source domains and bonding states continually transform and reconfigure. Narratively, this instability and ongoing

transformation sonify the AI system's training process, in which materials are continually reinterpreted and reorganised.

Output (7'28–10'07)

The output phase presents a fragmented display of AI-generated sounds (Groups *b* and *c*). This phase begins with erratic appearances of sound materials from Groups *b* and *c*, which are soon followed by more organised and periodic spectral activity, with different sound types recurring at relatively regular temporal intervals. In the closing section, a convolutional motion formed by low-band *c2* industrial friction and *b1* electrical hum produces a stable, forward-driving trajectory toward the work's last arrival.

This phase is defined by the dominance of AI-generated materials and by a shift from erratic behaviour to periodic organisation. Narratively, it represents the output stage of the system, in which, after a brief period of instability, the AI-generated sounds stabilise into a coherent and organised sonic result.

Overall, *Aura Machine* articulates its form through the organisation of sound behaviours and their distribution over time, while narrating the AI system's Input–Training–Output process as a trajectory from stability, through instability, to a newly constituted stability.

4. Discussion of AI Error Sounds in *Aura Machine*

Based on the detailed analyses above, this section addresses the questions posed in the introduction concerning how AI error sounds behave as compositional materials and what structural functions they perform within an electroacoustic work, and thereby leads into the concluding discussion of the aesthetics of AI error sounds in *Aura Machine*.

In this work, the behaviours of AI error sounds as compositional materials are inseparable from the structural functions they perform. Different from source sounds and AI source-bonded sounds, AI error sounds are not used to establish stable texture-setting fields or sustain continuity. Instead, most of them acquire the capacity to drive temporal progression through a set of sonic behaviours.

AI error sounds are often characterised by processes of agglomeration and endogeny. In *c1*, high-band micro-noise initially emerged as a texture-carried presence, and accelerated granulated motion develops. Through agglomeration and endogenic thickening, it leads to the low-band *a3* or *c3* impacts. Similarly, the *c2* industrial friction sounds evolve from pipe

friction to a granular sound body, while *c3* shifts from gesture-carried impact in the earlier stages of the work to gesture-framed sounds with a clear directional force later.

Through these characteristic behaviours, AI error sounds undertake a significant structural role by generating forward-driving pressure. Figure 6 compares the temporal proportions of AI error sounds relative to other sound types and, together with Figure 5, highlights their function in pushing such structural progress.

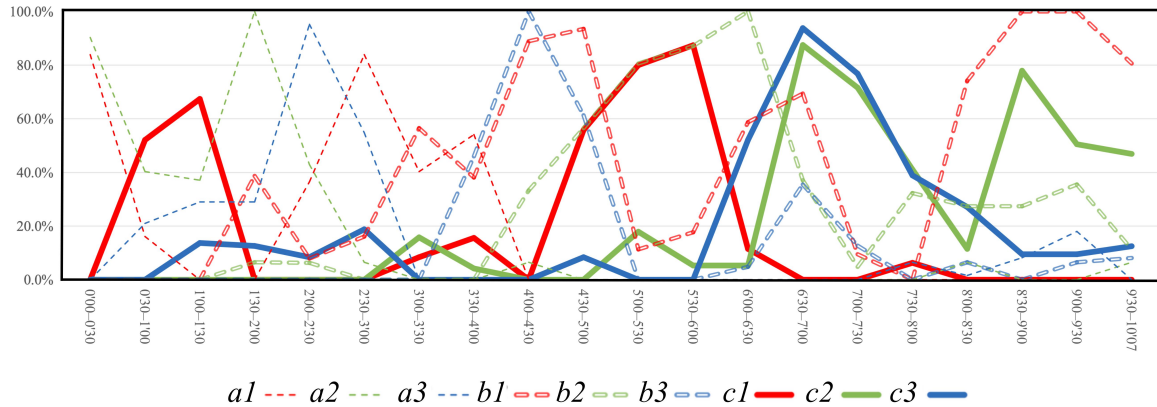


Figure 6. Temporal proportions of classified sounds calculated in 30-second intervals.

Throughout the work, AI error sounds are not evenly distributed. Instead, they gradually assume a more prominent position in sections characterised by enhanced transitions and instability, especially during the training stage, where they tend to cluster around points of structural change. By compressing spectral space and destabilising maintenance, AI error sounds drive the work forward and shape its structural progression. Compared with the source sounds that dominate the maintenance and texture-setting state and the AI source-bonded sounds that are often used to mediate gradual transitions, AI error sounds are more frequently associated with interruption, transition and arrival. This pattern indicates that their primary structural function lies in shaping the work's forward motion and constructing its overall form and narrative by controlling when and how the music progresses.

Overall, AI error sounds mostly push structural progression through agglomeration, endogeny, and disruptive transitions that compress spectral space and destabilise states of maintenance. Taken together, these behaviours form the basis of the aesthetic argument below.

5. Conclusion

In this essay, AI error sounds are not analysed through their underlying algorithms or technical identity. Instead, source bonding is employed as a perceptual bridge through which their spectromorphological identity is gradually established and consistently used as the primary framework for understanding AI error sounds in the analysis. Building on the preceding in-depth analysis, this essay therefore argues that the AI error sounds in *Aura Machine* acquire greater spectromorphological agency once their source bonding is substantially weakened.

Therefore, the analysis addresses a clear gap in the existing electroacoustic analytical research. Although established theoretical frameworks support treating technologically generated sounds as perceptible and analysable objects, and contemporary electroacoustic practices have increasingly incorporated AI error sounds as compositional materials, analyses of such sounds within specific works remain limited. By examining how AI error sounds behave and function as spectromorphological materials in *Aura Machine*, this essay contributes to a practical spectromorphological understanding of AI error sounds in the context of electroacoustic music.

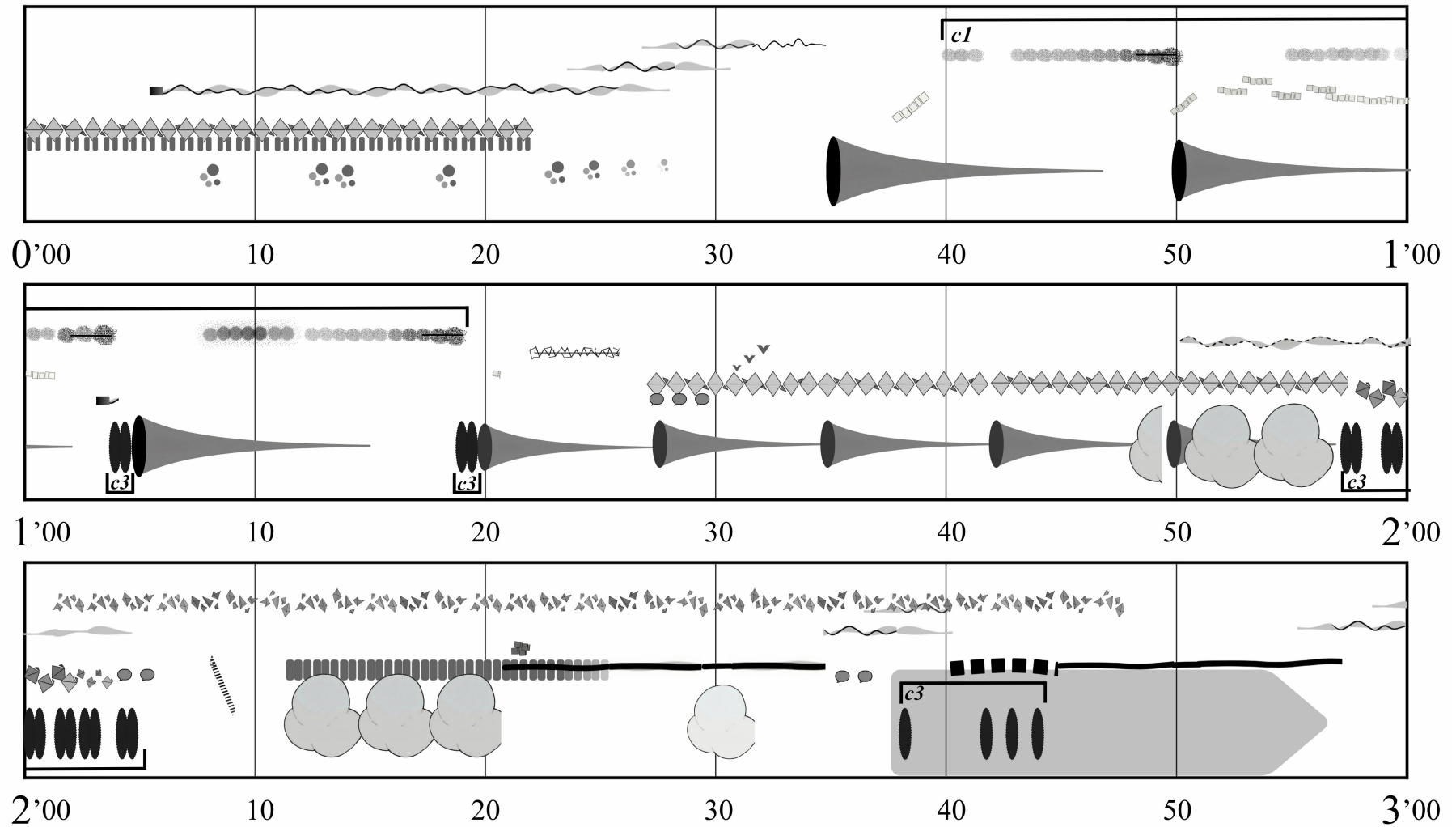
Bibliography

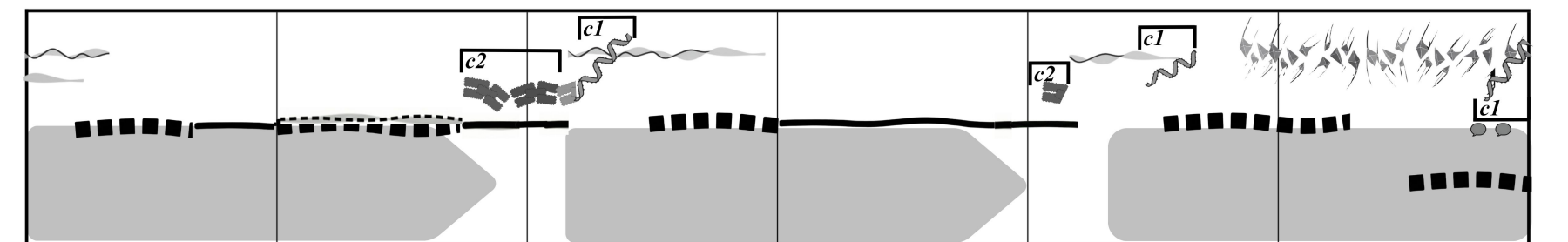
- Blackburn, Manuella, ‘The Visual Sound-Shapes of Spectromorphology: an illustrative guide to composition’, *Organised Sound*, 16 (2011), 1: 5–13.
- Blazsek, András, ‘Machine Listening as Sonification’, *Organised Sound*, 29 (2024), 3: 275–284.
- Caillon, Antoine, and Philippe Esling, ‘RAVE: A Variational Autoencoder for Fast and High-Quality Neural Audio Synthesis’, arXiv preprint, arXiv:2111.05011 (2021).
- Clarke, Vicky, ‘About’, Vicky Clarke, <https://vickyclarke.org/about/> (accessed 7 December 2025).
- Clarke, Vicky, ‘AURA MACHINE’, AURA MACHINE, <https://auramachine.blog/> (accessed 7 December 2025).
- Clarke, Vicky, ‘AURA MACHINE’, RNCM PRISM Blog, <https://www.rncm.ac.uk/research/research-activity/research-centres-rncm/prism/prism-blog/aura-machine/> (accessed 4 December 2025).
- Lewis, Andrew, ‘Francis Dhomont’s Novars’, *Journal of New Music Research*, 27 (1998), 1–2: 67–83.
- Shang, Zengqiang, Haozhe Zhang, Pengyuan Zhang, Linyi Wang and Ta Li, ‘Analysis and Solution to Aliasing Artifacts in Neural Waveform Generation Models’, *Applied Acoustics*, 203 (2023), 109183.
- Smalley, Denis, ‘Spectromorphology: Explaining Sound-Shapes’, *Organised Sound*, 2 (1997), 2: 107–126.

Appendix

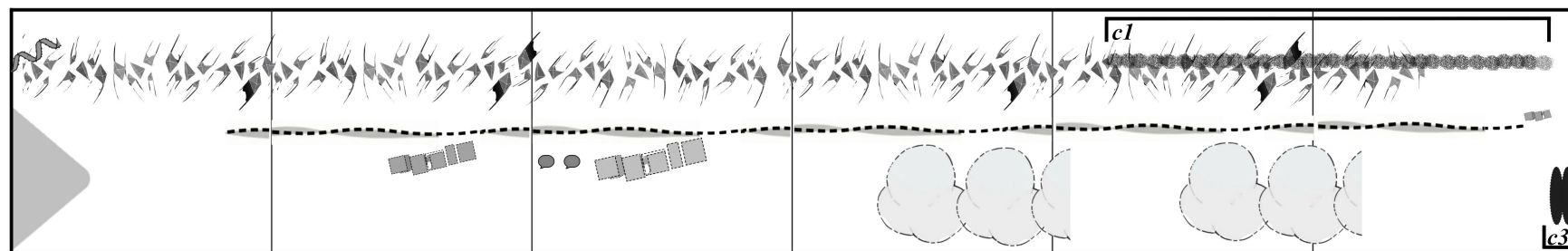
Aura Machine by Vicky Clarke

Scored by Tianrui Xiong

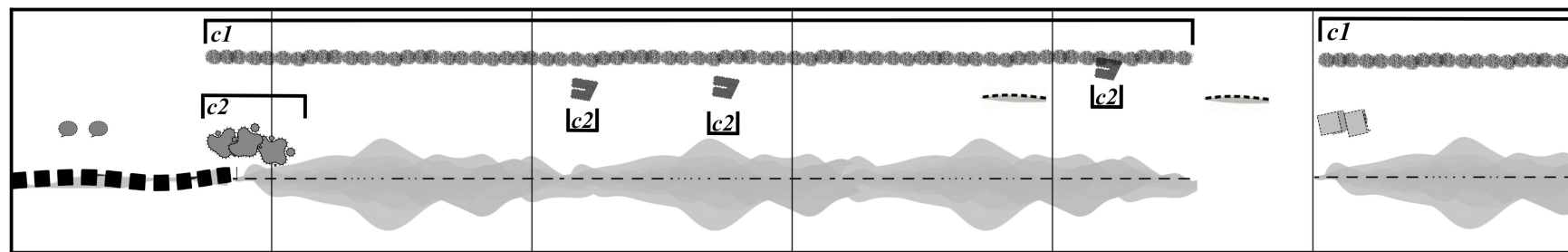




3'00 10 20 30 40 50 4'00



4'00 10 20 30 40 50 5'00



5'00 10 20 30 40 50 6'00

