

Latent Music: Simulacral Sound Objects and Interreferential Drifts

Guilherme Coelho

Audio Communication Department
Technische Universität Berlin
guilherme.coelho@thewnewcentre.org

Abstract

This installation explores what I call *latent music*: AI text-to-audio generated sound objects that remain in-becoming and occupy liminal sign-spaces between instruments and genres, characterised by spectral semiotics and transductive individuation in metastable fields. Instead of pursuing genre fidelity or stylistic imitation, the work deliberately employs prompt constructions that cultivate semiotic tensions and associative drifts, resulting in sonic artefacts that resist categorical stability and fixed form. The artefacts presented were catalysed through constructed word-assemblages functioning as floating signifiers—linguistic constellations designed to activate processes of cross-material association, reterritorialisation, and liminality within latent space, where encoded stylistic fragments are dislodged from their conventional positions and recombined into distinct configurations that exist at the threshold of stable recognition. Prompts and resultant audio are presented together: language scrolls as a mode of sonic world-making, rendering audible an intersemiotic translation in which words catalyse sound. The installation features selected latent-music objects paired with their generative prompts (semiotic scores), creating a dialogue between linguistic activation and sonic materialisation. Visually, the prompts are displayed in real-time, revealing the semiotic scaffolding that produces these reterritorialized sound objects. As a focused study of contemporary TTA systems, realised through Udio, the installation demonstrates how such models can generate singular, threshold-dwelling sonic forms.

Keywords: Latent music; Text-to-audio (TTA); Sonic Liminality; Spectral Semiotics; Intersemiotic Musicking; Prompt Assemblages; Semiotic Scores; Reterritorialization; Simulacral Sound Objects;

Exhibition Description

The term *latent* derives from the Latin *latens*, meaning “hidden,” “concealed,” or “lying in wait” (*Oxford English Dictionary*, n.d.). It denotes something that exists but remains unmanifest, awaiting the appropriate conditions for emergence. Across disciplinary contexts—from psychoanalysis and information theory to biology and philosophy—latency names a state of suspended potential: something present but not manifest, structurally embedded but not yet actualised. In psychology, latent refers to underlying, hidden, or unexpressed processes; in latent learning, knowledge is acquired without obvious motivation or reinforcement and becomes apparent only when there is a reason to demonstrate it. In machine learning, latent space designates a compressed relational manifold learned by the model, where proximities encode statistical regularities and along which prompts guide traversal. In philosophy, latency resonates with notions of virtuality (Deleuze, 1994), individuation (Simondon, 2005), and trace (Derrida, 1976). For Deleuze, the virtual constitutes the real as potential—the field of pre-individual multiplicities from which actual entities emerge. In this sense, the latent operates as a differential field of possibilities, structurally present within systems of becoming. Simondon’s concept of individuation similarly emphasizes *becoming* over being,

situating forms as emergent from metastable systems rather than pre-given structures of fixed form. Latency thus names a reservoir of incipient form—a differential field of possibilities, structurally present yet only actualised through indeterminate relations of becoming.¹

Drawing on a definition of “latent music” developed by accompanying research, this installation presents a selection of text-to-audio (TTA) generated sound objects characterised by liminality, spectral semiotics,² transductive individuation. I have investigated and generated thousands of outputs with Udio, and within this corpus certain recurrent behaviours emerge—interstitial timbral couplings, recombinatory reference, and non-teleological form. Their formation through latent-space traversal confers a haecceity—a sonic ‘thisness’—that would be difficult to realise in other media or production paradigms. These objects hover at the threshold of recognisability, intimate genres they never fully enter, and evolve through smooth yet unstable recombinations. Composed via prompt assemblages that braid signifiers from dispersed musical cultures, production idioms, and instrument ecologies, they rearticulate those traces within latent space, yielding distinctive objects of reterritorialisation. As a focused study of contemporary TTA systems, realised through Udio, the installation demonstrates how such models can generate singular, threshold-dwelling sonic forms.

In this sense, the work presents a study of how contemporary TTA models can generate distinct latent sound objects that navigate, blend, and cross-pollinate diverse signifiers and forms of representation. In concrete terms, some pieces read as plunderphonic in an analytic and aesthetic sense: an intertextual montage in which stylistic codes, production tropes, and genre-indexing signifiers are deconstructed and re-sutured by the model, yielding probabilistic reconstructions rather than direct quotation—collage by inference that merges coexisting materials drawn from disparate archives and scenes.³ Elsewhere, timbres inhabit interstitial spaces between instrumental signifiers: mallet harmonics might tilt between mbira and marimba; Auto-Tune-inflected voices might pivot between hyperpop sheen and synthetic plainchant, grazing pop’s hook grammar without settling into it; metallic decays might drift from cymbal to gamelan. Across these vectors, objects reterritorialise as they drift across musical territories.

Within these sound objects, a dense web of spectral traces surfaces—echoes of countless musical works, production techniques, and sonic signatures compressed into the model’s latent space. These emergent simulacral presences simultaneously haunt and constitute the outputs without fully materialising; their organisation is rhizomatic, a becoming-with of fragments that recombine along

¹ For instance, in text-to-audio systems such as Udio, the same prompt will almost never yield an identical output across multiple generations. This is due to both the stochastic nature of the generation process and the model’s architecture, which allows for a range of interpolations within latent space. Each prompt thus carries a verisimilitude of potentialities—a probabilistic envelope of sonic outcomes shaped by semantic proximity, stylistic bias, and the specificity or ambiguity of the prompt itself. The more distinct or semantically ambiguous a prompt is—particularly when composed of unconventional or unrelated signifiers—the more likely it is to activate divergent interpretative paths within the model, producing outputs that are varied, unstable, and singular.

² I use spectral semiotics to denote a framework for reading signs and ideologies that are present yet elusive—ghost-like traces that circulate without a stable origin, often within globalised and platformed media ecologies. Meaning emerges through corrupted, fragmented, or displaced signifiers, producing a “spectral” effect in which source and definitive reference are deferred or undecidable (Derrida, 1994; Gordon, 2008; Fisher, 2014; Blanco & Peeren, 2013).

³ **Audio Examples.** These initial examples demonstrate the characteristic liminality and interreferential drift of latent music. Listeners are invited to observe the unstable timbral identities and how musical signifiers fluctuate across recognizable territories without resolving into fixed genres or instrumental categories. The resulting sonic artifacts exist in states of perpetual becoming—simultaneously familiar and estranged, referential yet resisting complete categorization. Link to audio repository: <https://drive.google.com/drive/folders/1KfihFG9jkXDz56pdg4xhYfGXDd-MVRGC?usp=sharing>

non-hierarchical paths. The pieces thus occupy a zone where many works are present and absent at once, referenced yet transformed: a perpetual becoming in which the boundaries between original and derivative, presence and absence, are fundamentally destabilised. These are simulacral formations that evoke without reproducing and reference without resolving—manifesting liminal semiotics and gradient identities distributed across continuous spectra and porous categorical borders. Once heard, the effect is unmistakable: like tuning a radio while the stations bleed into one another, or reading a sonic palimpsest where erased lines keep returning—identity flickers just long enough to be felt before sliding sideways. Coherence arrives, then slips away.

Within text-to-audio systems, activation occurs through prompts—textual sign-units that act as semantic catalysts in the model’s latent manifold. These prompts function as assemblages of words carrying stylistic, cultural, and affective associations that guide traversal across learned proximities. A phrase such as “avant-garde, plunderphonics, folk, glitch music” does not resolve into a single genre or form; it initiates a vectorial search through multiple regions at once. Each term lights up clusters of embedded sonic features, and their conjunction yields a recombinant field of potentials. In this sense, prompts operate as *semiotic scores*; language functions as vectorial scaffolding for latent traversal and object formation, so meaning arises from relations among sign-units and their associations in the topology rather than any single token. Because the manifold is continuous, such traversals reach interstitial sign-spaces where instrumental and generic markers overlap or blur.⁴ TTA models are uniquely suited to this exploration: by interpolating across learned regularities, they surface mixtures that challenge categorical listening, inhabit the interstices between signifiers, and facilitate the reterritorialisation of familiar idioms into distinct forms.

The installation makes this intersemiotic relay explicit. Each sound object is accompanied by its generating prompt, unfurled as a textual projection. As words appear and recombine, the audience interprets the semiotic scaffolding that conditions what becomes audible. The prompt–audio pairing discloses a double movement—linguistic activation and sonic materialisation—in which encoded stylistic fragments are dislodged, recombined, and reterritorialised into configurations that remain at the threshold of recognition. This coupling enacts intersemiotic musicking and schematic association, inviting listeners to trace how human schemas and model embeddings co-produce meaning. The prompt can be heard as a procedural semiotic score; the sound as the manifold’s response—a negotiation between linguistic cues, learned representations, and situated listening.

Crucially, *Latent Music* invites a corresponding mode of reception. Listening is treated as speculative and associative rather than classificatory: a practice of dwelling with unfinished forms, tracking drifts and inflections across perceptual thresholds, and attending to how meaning is inferred rather than given. Listeners are invited to hear these works as digital entities with extensible, non-finite ontology—nodes in open chains of continuation—aligning musical experience with media-theoretical accounts of TTA artefacts as processual and recombinable.

The work foregrounds five recurrent properties of latent music: (1) liminal identities that refuse categorical settlement; (2) timbral hybridisation that blurs acoustic/electronic/synthetic provenance; (3) interreferential drift, where cultural and stylistic traces circulate as spectres rather than quotes; (4) formal indeterminacy, privileging emergence over teleology; and (5) recombinatory semiotics, in which textual sign-assemblages and deconstructed embeddings orchestrate audio indeterminately. The result sounds like latency itself—adrift, recombinant, and perpetually in the act of becoming.

⁴ The model performs a form of associative synthesis—navigating across encoded sonic materials and generating outputs that reflect the relational tensions and affinities inscribed in the prompt’s internal structure.

Technical Rider

Contributor

Guilherme Coelho is a lecturer and doctoral researcher exploring the entanglements of music, technology, and media theory, with a particular emphasis on AI and text-to-audio paradigms. Their dissertation, *Text-to-Audio AI: Technological Architectures, Mediating Processes, Latent Exploration, and User Interaction*, offers the first comprehensive scholarly study of text-to-audio systems, integrating aesthetic, intersemiotic, and media-theoretical perspectives. Drawing from medium theory and sonic materialism, their work critically interrogates how tools shape cultural production, sonic identity, and creative agency. They have designed and taught courses on AI and music, examining shifting modes of musicianship, representation, and mediation. Their academic contributions are complemented by industry experience in music production and sound design, with credits including Korn's *The Nothing* album and commercial collaborations with Axel Arigato, Rimowa, and Kölner Haie. Their practice spans critical theory and speculative sonic creation, situating AI-generated sound within broader discourses of digital aesthetics and contemporary artistic practice.

Technical Equipment Provided

- Laptop with sound files and visual presentation
- USB drive with backup of all installation materials

Setup Requirements

- Stereo amplification (provided by venue)
- Video projection capability for displaying prompts (projector or large display)
- One power outlet
- Connection to house sound system (1/4" TRS or XLR)
- HDMI connection for visual projection
- Setup time required: 5 minutes

Installation Structure: The 20-minute cyclical installation consists of several sound objects, each preceded by its generative prompt displayed visually. The visual component displays the prompts as they unfold, revealing the linguistic structures that give rise to these sonic artifacts.

Documentation: Examples of the sound objects to be included in the installation can be accessed here:

https://drive.google.com/drive/folders/1tpwCsVv7HIG1ixzK_nYyFCGnD0NEtgiF?usp=sharing

Accessibility Requirements: None applicable

References

- Blanco, M. del P., & Peeren, E. (Eds.). (2013). *The Spectralities Reader: Ghosts and Haunting in Contemporary Cultural Theory*. Bloomsbury.
- Deleuze, G. (1994). *Difference and Repetition* (P. Patton, Trans.). Columbia University Press. (Original work published 1968)
- Derrida, J. (1976). *Of Grammatology* (G. C. Spivak, Trans.). Johns Hopkins University Press.
- Derrida, J. (1994). *Specters of Marx: The State of the Debt, the Work of Mourning and the New International* (P. Kamuf, Trans.). Routledge.
- Fisher, M. (2014). *Ghosts of My Life: Writings on Depression, Hauntology and Lost Futures*. Zero Books.
- Gordon, A. F. (2008). *Ghostly Matters: Haunting and the Sociological Imagination* (2nd ed.). University of Minnesota Press.
- Oxford English Dictionary. (n.d.). *Latent, adj. & n.* In *OED Online*. Oxford University Press.
- Simondon, G. (2005). *L'individuation à la lumière des notions de forme et d'information*. Jérôme Millon. (Original work published 1958)