Generative audio for music-making (GAMMa) workshop 2026

December 4 décembre, 2026, 09:00 – 16:00
A-832/A-833 (8 floor/étage) Elizabeth Wirth Music Building, 527 Sherbrooke O

Workshop Schedule with Abstracts / Programme de l'atelier avec résumés

09:00 — *Welcome/Bienvenue*

Session 1: Case studies / Études de cas

09:20 — James Ford Stinchcombe [in-person/en personne]: *Mobile instruments for gestural navigation of latent spaces*

I will discuss my findings of navigating latent spaces Using a mobile phone passively augmented with 3d printing. (What my gestures reveal to me about the latent space, and how latent spaces augment my performance practice.)

09:40 — Zi Hang Yin [in-person/en personne]: *Two Voices of Gesture: A Dual-Microphone Interface for Generative Audio Performance*

This research-creation project proposes a dual-microphone interface for live interaction with generative audio systems. One microphone captures vocal gestures, including speech, breath, and nonverbal sounds, while a second captures hand-produced sounds such as tapping, clapping, and friction. The two channels serve complementary roles: vocal features shape timbral characteristics, while hand gestures influence rhythmic articulation and event triggering.

Building on recent research into navigable timbre spaces and generative musical instruments, particularly Thibault and Piazza's WÆnderer, the project explores how ordinary acoustic actions can become expressive controls for neural audio synthesis.

The presentation will examine the design of this multimodal mapping, its potential for embodied musical interaction, and the challenges of latency, controllability, and acoustic interference. A work-in-progress demonstration is envisaged, with particular attention to the creative possibilities and limitations of performing with two independent acoustic input streams.

10:00 — Nicola Casetta [online/en ligne]: *Variation Is Not Memory: Two Feedback Instruments Built Around a Frozen RAVE Model*

Eigenform couples a closed nonlinear feedback network with a RAVE model trained on about three hours of recordings of its predecessor: a small, bespoke dataset turned into a fixed memory of earlier sound. Ecphory, a work in progress, adds a second, adaptive memory: an online Gaussian mixture in RAVE's latent space that steers what the frozen decoder plays and learns from what the instrument and the room return. When it learns only from its own output, this becomes a self-consuming loop. Offline experiments revealed an unexpected outcome: the loop does not collapse, but what keeps it open is the encoder's random sampling, noise of constant size that carries no information about the input. Variance survives; slow structure, diversity and level do not. The talk presents both instruments with audio examples and discusses what this suggests for the next generation of latent-space instruments: encoders whose variation depends on what they hear.

10:20 — Vincent Isabel [in-person/en personne]: *Interfacing audio generative AI models with DIY noise machines*

The presentation will show the work in progress of the sound-making systems I am currently building, where I am experimenting how DIY digital instruments can interface with audio generative artificial intelligence models.

10:40 — *Coffee Break/Pause Café*

Session 2: Philosophy & practices / Philosophie et pratiques

11:00 – Mikael Meunier-Bisson; Jacob Desjardins [in-person/en personne]: *Les machines sont-elles des Êtres ? Agentivité, milieu sonore et production des subjectivités au temps de l'audio génératif*

Le courant du tournant ontologique mobilise la notion d'agentivité pour défaire les hiérarchies sociales (formes de vies humaines/non-humaines). Le terme s'immisce en art (sonore) de manière consensuelle, mais pourquoi? Que cherche-t-il à produire? À la fois comme concept analytique et opératoire, nous souhaitons interroger la nécessité de l'usage du terme pour penser le rôle des outils technologiques, plus précisément notre relation aux dispositifs de génération audio. Nous croyons que d'autres concepts seraient plus féconds pour décrire la relation, notamment celui de 'milieu sonore'. De manière exploratoire nous proposons de réfléchir : le processus continu qui produit les « sujets » non-fixes; le discours du dépassement humain (transhumaniste/posthumaniste) encouragé par l'usage du terme; et les possibles qu'une bifurcation conceptuelle adaptée pourrait ouvrir. La présentation durera 15-20 minutes, suivie d'une période d'échange avec les participant-es et (possiblement) d'une démonstration de patch génératif de synthèse par feedback.

11:20 – Misagh Azimi [in-person/en personne]: *The work of music in the age of commercial GenAI: Politicisation, prophecy, and participation*

Commercial GenAI music platforms promise democratic creation, yet their operating logic—what I term fragmentation and reconstruction—fundamentally alters the social, aesthetic, and economic conditions of music. Rather than reproducing or distributing existing works, these systems fragment millions of recordings into statistical distributions and reconstruct new outputs from learned patterns, severing the traceable link between a musical artefact and the identifiable human labour that produced it.

On the sociocultural side, I argue that this paradigm grants the creator and the audience expression without agency; writes the musicians whose work built these systems out of the relationships that work sustained, and locks music into repetition at a scale that forecloses the kind of genuine break that has historically pushed musical culture forward. On the economic side, I show that these platforms combine large-scale cultural extraction, rent-seeking over digital infrastructure, and a new form of locked labour in which musicians continuously generate value while being structurally excluded from its distribution. Crucially, however, I close by turning toward repoliticisation of music: through interventionist and practice-based work, including my own, I demonstrate that musicians can reclaim agency, restore traceable creative relationships, and steer generative AI systems from instruments of extraction toward instruments of collective critique.

11:40 – Gustavo Guzmán [online/en ligne]: *A Capricious Translation, or On the Limits of Mapping Gesture Across Domains*

This presentation traces gesture in electronic music from the Theremin onward, asking how bodies become signals, signals become parameters, and mappings become musical thought. It places these translations in conversation with other arts and with shifting ideas of interaction—from Subotnick's control-voltage contours to contemporary systems for cross-modal parameter mapping. Gesture, in this view, is less a command than a translation: shaped by subjective appropriation, material affordances, contingent infrastructures, and the frictions that make a system playable, expressive, or resistant. These are conditions through which artistic practice and technical invention take shape in tandem. The discussion then turns to neural networks and latent spaces as explicit systems of mapping, asking what they make actionable and what they obscure in musical creation. With historical perspective and a set of unresolved questions, the presentation considers what generative instruments might learn from gesture—and what gesture refuses to translate without loss.

12:00 – Kasey Pocius [en personne]: *From 3tube to the Edge and Back*

This presentation traces my work as a composer, performer, and technologist with the eTube project over several years, focusing on the artistic and technical challenges of composing, performing, and re-performing works centred on real-time interactions between performers and machine-learning models. Particular attention will be given to the evolving relationships between performer, instrument, and computational processes, alongside questions of reproducibility, adaptability, and preserving interactive works over time.

These experiences will be contrasted with my parallel compositional practice, including fixed-media and solo performance works. By examining the intersections between these approaches, the presentation explores how techniques developed through group interactive performance have informed my solo compositional output, and how these practices have, in turn, influenced my approach to designing and performing with interactive systems in group settings.

12:20 – 13:20: Lunch / Dîner

Session 3: Symbolic generation / Génération symbolique

13:20 – Pierre Basso [in-person/en personne]: *CAMBIATA : générer du contrepoint d'espèces avec les grands modèles de langage*

CAMBIATA est un corpus de contrepoints d'espèces générés par trois grands modèles de langage (Claude Opus 4.6, GPT-5.4 et Gemini 3.1 Pro). À partir du Gradus ad Parnassum de Fux (1725), nous avons rédigé un guide d'instructions pour chaque espèce, puis obtenu par API 450 contrepoints en MusicXML. Un évaluateur automatique (Python, music21), fondé sur une taxonomie de 71 règles et complété par une lecture experte, relève les fautes de chaque contrepoint. Les résultats montrent des écarts nets entre les modèles (de 3,37 à 8,92 fautes en moyenne), des difficultés persistantes dans le traitement de la dissonance et une forte tendance à la répétition, puisque 31 % des contrepoints sont identiques note pour note à un autre. Le travail est symbolique plutôt qu'audio, mais il interroge ce que ces modèles saisissent réellement des contraintes musicales. Nous présenterons l'affiche primée à EuroMAC 11, avec des exemples commentés.

13:40 – Uandha Fernandes Barbosa [in-person/en personne]: *Computer Say No*

This project begins with reflections on AI's structural inability to say "no." Milgram asserted that obedience was a determining behaviour relevant in his time, which remains considerably current in our world, and might include our AI language-driven tools. Computer Say No re(creates) shock and interview rooms with exaggerated silicon personas, bringing LLM's into the 60s experiment to see what causes them to "(dis)obey" and to ask: what does this refusal mean if it, too, is engineered?

The project is now at its first-stage sonification experience. Using live coding in TidalCycles with a implemented syntax, the personas' words and phrases are translated into Rorschach-like figures, whose shapes become trajectories for navigating sound corpora in AudioStellar. The corpora are being curated and are still under development to also express the personas' responses, emotions, and reflections—sometimes deeply unsettling.
Category: ongoing · live coding corpus navigation · concatenative · LLM personas

14:00 – Nicholas McCullough [in-person/en personne]: *Developing Generative Music Systems using Tone Clock Theory for Jazz Composition, Improvisation, and Performance*

This presentation explores a generative performance system built in Max/MSP that applies Peter Schat's and Jenny McLeod's Tone Clock Theory to electroacoustic jazz improvisation, composition and live performance. The system pairs a probability based sample sequencer with a generative synthesis engine whose pitch material is drawn from tone clock "hour tonalities" (twelve-note rows) over a randomized root. Transitions and dwell times set by a human performer between triadic shapes turn randomization into harmonic motion and performance shape rather than noise, while minimalist controls keep the system playable with limited dexterity. Drawing on my use of systems with a physical interface, hardware effects, and acoustic saxophone during a November-December 2025 European tour and McGill's 2026 Research Alive colloquium with Jonathan Lindhorst, I will discuss developing it as an improvising partner where human control meets algorithmic chance. Future work includes gestural control. The talk includes a demonstration and invites discussion of accessible, performable generative instruments for various performance disciplines.

14:20 – David Piazza [in-person/en personne]: *Experiments in Code Synthesis and Small-Data Neural Models for Acousmatic Music Composition*

This presentation reports on a work-in-progress; an 8-channel fixed-media piece composed entirely from a bank of 50 recorded sounds, and on the two generative practices that shaped it. The first is code synthesis: I describe gestural and spectral characteristics to a coding agent, which writes batch-processing scripts that render sound units with timbral and spatial trajectories already inscribed in them. The second is a set of small DDSP models, derived from the architecture of NoiseBandNet, trained locally on each instrumental category of the same 50 sounds. The piece attempts to build on a dialectic between these transformed source sounds and their neural simulacra. I will play short excerpts and discuss what my workflow affords in practice: where text description succeeds or fails as an interface to sound, what models trained on very little audio can and cannot learn, and what questions about authorship and provenance remain open.

14:40 — *Coffee Break/Pause Café*

Session 4: Musical agents / Agents musicaux

15:00 – Keon Ju Maverick Lee [in-person/en personne]: *Beyond Prompt-to-Product: Artist-in-the-Loop Musical Agents, Live Improvisation, and Cooperative Futures for Music AI*

Much discussion of music AI focuses on generating finished songs. This presentation shifts attention to live performance, where musical agents must listen, respond in time, negotiate initiative, and recover from unexpected behaviour. Drawing on interactive music systems, musical metacreation, and research-creation examples involving audio, MIDI, embodied, and audiovisual agents, I outline six design principles for artist-in-the-loop co-creation. I then propose a multi-level approach to evaluation that considers system behaviour, performer experience, audience understanding, and cultural consequences. A five-stage cultural hegemony transformation framework provides a secondary lens for examining how musical AI gains legitimacy without assuming that acceptance equals justice. Finally, I distinguish interactional agency—the ability to shape an agent's musical behaviour—from infrastructural agency—the ability to influence data, platforms, and governance. The presentation considers cooperative and federated futures in which artists help shape both musical technologies and the institutions surrounding them. The emphasis is on artistic practice rather than replacement.

15:20 – Vincent Cusson [in-person/en personne]: *How Many Is an Agent? Scaling Generative Co-Improvisation to Many Voices and Many Players*

Most generative music systems are built around a duo: one performer, one model. This presentation reflects on what changes when both sides of that relationship multiply. Drawing on my work on a distributed vocal installation developed with IRCAM's ISMM team and presented in France, I discuss two scaling moves. The first goes from one agent to more than a hundred corpus-based agents, each drawing on a small sound corpus. At that scale, the boundary of "an agent" blurs: the system can be heard as a crowd of small voices or as one organism made of micro-behaviours. The second goes from one trained performer to several untrained visitors singing at once. Density piles up, the link between gesture and response dissolves, and legibility becomes the main design constraint. I share the strategies we tried, what participants reported, and open questions for generative systems designed for groups rather than individuals.

15:40 – Dominic Thibault [in-person/en personne]: *WÆnderer: Projecting and Navigating User-Defined Audio Corpora in Pre-Trained Audio Models*

Neural audio synthesis offers powerful ways to create sound, but many systems remain difficult to use as musical instruments. Artist-specific models demand large, curated datasets and training resources, while accessible text-to-audio services give limited control over personal material and often obscure data provenance. We present WÆnderer, a system for real-time sound generation that repurposes a pre-trained variational autoencoder (VAE) - a model that compresses audio into a continuous latent representation and reconstructs it through a decoder - without retraining the audio model. WÆnderer encodes small, artist-defined corpora, organizes their sound segments in an abstract timbre space derived from audio descriptors, and lets performers navigate that space while the VAE reconstructs retrieved latent frames. The system therefore does not generate audio from text or unconstrained latent sampling; it reconfigures a musician's own recorded material within the representational limits of the pre-trained model. A practice-led evaluation identifies three design findings: corpus-aligned retrieval avoids artifacts observed during direct interpolation; decoding-window length trades responsiveness against continuity; and manual, trajectory-based, and morphology-based navigation produce distinct forms of control and sonic organization. This research-creation study contributes an accessible, locally executable approach to neural audio interaction and frames latent-space performance as navigable, corpus-based music rather than autonomous generation.