The Hard Problem of Text-to-Music: Structural Parallels with Consciousness

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

October 23, 2025

Text-to-music generation models have made remarkable progress in capturing mood, timbre, and genre from natural language prompts. Yet even the most advanced systems—whether trained on symbolic scores or raw audio—struggle with coherent development across time. They can evoke a feeling, but rarely build a theme; they sound musical, but seldom narrate. This failure of thematic integration is not merely a technical hurdle—it mirrors the "hard problem" of consciousness: how discrete, local signals give rise to unified, structured experience. This paper explores the deep structural parallels between these domains. In both, a sparse or fragmented input (text, neural signals) attempts to generate an emergent output (music, experience) with rich internal coherence. The challenge lies not in style, function, or response—but in form, memory, and development. We examine how motifs in music resemble qualia in consciousness, how musical form mirrors unity of experience, and how breakthroughs may require entropic tunneling—discontinuous shifts in representational grammar. Ultimately, this analysis reframes text-to-music generation as a testbed for studying emergence, binding, and structure—not just in sound, but in minds.

Keywords: text-to-music, consciousness, emergence, motif, qualia, entropic tunneling, thematic development, coherence, hard problem, unity, experience, structure, musical form, generative AI, perception, q to τ.

1. Introduction

1.1 The Promise and Limits of Text-to-Music Systems

In recent years, large-scale models like MusicLM, MusicGen, Jukebox, and Suno have made text-to-music generation a commercial and creative frontier. These systems can convincingly synthesize short musical segments that reflect the mood, genre, and instrumentation implied by natural language prompts. However, their coherence often breaks down beyond a few dozen seconds. Thematic development stalls, motifs are rarely recalled or transformed, and long-range form lacks direction. Despite advancements in style and timbre control, these systems struggle with what musicians call *form*—the intentional unfolding of motifs across time. The failure here is not due to data scarcity or weak architectures, but rather due to a deeper representational mismatch between text and music. Words evoke; they do not compose.

1.2 The Analogy with Consciousness: Structure, Not Signal

This breakdown in text-to-music coherence closely resembles what philosophers call the hard problem of consciousness. Just as brain activity can be mapped and modeled without yielding first-person experience, text prompts can be transformed into musical textures without yielding form, meaning, or memory. In both cases, *signal is not enough*. What's missing is structure—the binding of elements across time into a unified experiential or thematic whole. Just as neurons firing do not explain the unity of consciousness, well-styled audio does not explain the unity of a musical work. The structural analogy is more than metaphorical—it may be formal.

1.3 Overview of Argument and Stakes

This paper argues that the limitations of current text-to-music systems and the mystery of consciousness share a common informational topology. Both involve a divergence between a low-dimensional control input (q) and a high-dimensional generative structure (τ) whose form is neither emergent from data alone nor reducible to local function. We develop this parallel through the lens of *motifs* (musical) and *qualia* (experiential), treating both as atomic units of sensation requiring development, binding, and recurrence to become meaningful. Drawing on principles from information geometry, cognitive theory, and musicology, we propose a constructive research program that reframes music generation as an epistemic testbed for emergence. The stakes are high: understanding how structure emerges from prompt, or experience from signal, could reshape both the future of AI creativity and the science of consciousness.

---

2. The Generative Gap: Underdetermination and Emergence

2.1 From Sparse Prompts to Structured Outputs

Text-to-music generation begins with a sparse prompt: a few words describing a mood, a genre, or a scene. From this limited input, the model must unfold a multi-layered structure across time—melody, harmony, rhythm, form, and expression. This transformation resembles zero-shot composition: from the sentence *"a sad song in the rain"*, a system must decide key, tempo, phrasing, motif selection, harmonic progression, and evolution across minutes of sound. The core challenge is this: *the prompt doesn't contain enough information to uniquely specify the output*. It opens a hyperspace of plausible interpretations—the model must not only select one but make it feel *intentional*.

This is the generative gap: the distance between low-dimensional guidance and high-dimensional structure. The gap cannot be closed by stylistic mimicry alone; what is needed is developmental logic—a policy for expanding, varying, and revisiting motifs in ways that satisfy both expectation and surprise. Without such logic, outputs may sound musical but lack form, just as neural spikes may support cognition but lack experience.

2.2 Why Language Underspecifies Music (and Brains Underspecify Minds)

Language is propositional, discrete, and referential. Music is continuous, recursive, and affective. A word like *"triumphant"* may evoke trumpets or a key change to major—but it doesn't encode melodic contour, counterpoint, or dynamic progression. The semantic bandwidth of language is orders of magnitude lower than the information content of even a few bars of music. This mismatch causes language-based generators to fixate on style and texture, while form and memory remain emergent—or worse, ignored.

This same underdetermination haunts neuroscience. A scan of brain activity during a visual task can suggest *where* something happens—but not *how* the subject experiences it. The neural signal is sufficient for behavior, but insufficient for phenomenology. Both cases point to the same asymmetry: a low-resolution controller trying to evoke a high-resolution output.

The common thread is structure. In both music and consciousness, coherence arises not from stimulus-response mappings but from temporal architectures—recall, variation, feedback, and binding. These are not *generated* by the prompt or signal—they are imposed *through* it by something like a formal grammar of becoming.

2.3 Entropic Bottlenecks in Both Domains

When a model generates music from a prompt, it initially samples a rich cloud of possibilities. But without structured constraints, the output often collapses into entropy—wandering arpeggios, flattened repetition, or style with no story. The same applies to cognitive simulations: when neural correlates are overfitted without deeper architectural grounding, we get function without form—behavioral mimicry without experience, simulacra without subjectivity.

We call this a kind of entropic bottleneck. It's not that the system lacks variation—it has too much. But without form-shaping constraints—like motif grammars in music or unity-of-consciousness principles in mind—the variation is not *funneled into emergence*. It disperses. Structure becomes accidental rather than causal.

Escaping this bottleneck requires discontinuous shifts in representation: reorganizing how motifs are selected and developed, or how experience is bound and recalled. These shifts are not smooth interpolations. They require entropic tunneling—a jump from one coherent schema to another. In music, this might look like a return of a transformed theme; in consciousness, it may feel like a flash of self-awareness. In both, it is structure that breaks the spell of entropy, not signal alone.

3. Motifs and Qualia: Atomic Units of Experience

3.1 What is a Motif? What is a Quale?

A motif in music is a short, recognizable pattern—of notes, rhythm, or harmony—that acts as a seed for larger structures. It is not merely a phrase; it is a *generative unit*. A motif gains meaning by being recalled, varied, juxtaposed, and transformed across time. Beethoven's Fifth Symphony famously begins with four notes: da-da-da-DA. That motif becomes a genetic code for the entire movement. In this sense, motifs are not just patterns—they are compressed intentions.

In consciousness studies, a quale (plural *qualia*) refers to the irreducible units of subjective experience—*the redness of red*, *the bitterness of coffee*, *the ache of nostalgia*. Just as motifs are the felt material of music, qualia are the felt material of experience. They are atomic not in scale, but in epistemic role: they are *what it is like* to be aware of something.

Both motifs and qualia present an unresolved mystery: how such *local events* come to participate in *global experience*. They are intelligible in isolation—but their significance unfolds only in a context of recurrence, contrast, and development.

3.2 Local Distinctness vs Global Integration

Isolated, both motifs and qualia are distinct and salient. A motif can be sharp, haunting, or joyful in its opening appearance. A quale—like the sensation of heat—is vivid, immediate, and undeniable. Yet neither achieves *meaning* without integration into a larger whole.

This tension between local distinctness and global integration defines the compositional challenge in both domains. In music, a motif must be echoed, inverted, delayed, layered—only then does it acquire narrative weight. Similarly, in consciousness, a quale must be bound to time, memory, and self to become part of the stream of awareness. Without integration, motifs are just sound-bites; qualia are momentary flickers.

This raises a structural problem: what mechanisms—or grammars—bind these local units into *coherent arcs*? In music, this is the role of thematic development. In consciousness, it may be played by binding dynamics, attentional coherence, or higher-order global workspaces. In both, the phenomenon to explain is *cohesion without collapse*—how variety forms unity without becoming monotony.

3.3 Theme as Memory: The Musical Binding Problem

To create a theme, a composer doesn't merely repeat a motif—they develop it. They treat it as memory does: revisiting, transforming, weaving it through new contexts. A rising third may return as a falling sixth; a rhythmic pattern may quicken, split, or dissolve. This is developmental memory, not rote repetition. It requires *intentional recall with modulation*—a direct parallel to how the mind recalls, mutates, and reinterprets qualia across time.

This gives rise to the musical binding problem: how do motifs retain identity through change? And how does a listener recognize a transformed version of an earlier idea as "the same"? The mind performs a kind of informational alignment—not of surface features, but of deep structure. This process is not unlike the persistence of self across qualia. You are the same "you" across radically different feelings and thoughts—not because your states are the same, but because their development is continuous.

Thus, theme is musical memory in the same way that self is experiential memory. Both require:

- Initial anchoring (motif or quale)
- Development through variation
- Contextual reinforcement (harmonic or narrative)
- A listener or subject to bind it into one story

The failure to develop motifs is not just a stylistic shortcoming; it is *a structural forgetfulness*—a musical analogue to amnesia or dissociation. A system that can generate notes without binding them into a theme may sound musical, but it will never sound intentional. Likewise, a system that simulates neural activity without experiential integration may act conscious, but will never *be* it.

4. Development and Unity: Form as Experience

4.1 Musical Form and Narrative Coherence

Musical form is not merely a container for sound—it is the *architecture of meaning*. Classical structures like sonata form, AABA, or verse-chorus-bridge do more than organize time; they guide expectation, reward memory, and shape emotional arcs. Even in minimalist or ambient styles, development remains critical: the return of a motif, the variation of rhythm, or the evolution of harmony signals a *narrative logic*. This is what separates structured composition from mere sound collage.

At its best, musical form creates a temporal unity—a felt progression where each part resonates with the whole. This mirrors narrative coherence in storytelling: the way an early image foreshadows a later turn, or how a resolution brings thematic closure. Crucially, this coherence is not imposed externally but emerges from development—from how motifs are bound together across time into a higher-order experience.

4.2 Consciousness as Unified Unfolding

Consciousness, like music, is a stream—not a series of snapshots. It unfolds as a structured continuity, where perceptions, memories, and emotions form a temporally extended unity. The philosophical challenge is not explaining *why a brain state exists*, but how disparate sensations are bound into a single unfolding awareness—what William James called "the stream of thought."

This unity cannot be reduced to simultaneity. Multiple stimuli may co-occur, but only *some* are experienced as coherent. Likewise, a musical piece may contain many sounds, but only some are thematically integrated. The analogy here is deep: in both cases, binding is not a function of signal density, but of development. It is *how* things relate across time—not *that* they co-occur—that creates unity.

Models of consciousness such as Global Workspace Theory or Integrated Information Theory (IIT) attempt to explain this binding, but often struggle with the explanatory gap between local

representation and felt experience. The same gap appears in music generation: we can specify local features (key, tempo, instrument), but without developmental logic, the result lacks intentionality. Both systems demand a theory of structural coherence over time—of unfolding that feels meaningful.

4.3 Failures of Development as "Zombie Outputs"

The philosophical zombie is a being that behaves exactly like a conscious person but lacks subjective experience. In the same way, a text-to-music model can generate sequences that *sound* musical—proper notes, pleasing textures, stylistic consistency—yet fail to develop. These are zombie outputs: *formally correct, emotionally void*.

Such outputs may pass brief listening tests, just as a chatbot may pass Turing-like exchanges. But sustained engagement reveals the emptiness: no arc, no transformation, no return. The motifs, like qualia in a zombie, appear but do not participate in anything higher. There is no *aboutness*, no evolving intentionality—just wandering signal.

This is not merely a stylistic critique. It is a deep epistemic failure. When systems cannot bind motifs into form—or sensations into self—they demonstrate the limits of local function without global coherence. Development, in this sense, is the litmus test for *true generativity*. It marks the difference between *output that exists* and *output that means*.

The challenge, then, is to design systems—musical or cognitive—that tunnel past entropy, binding atoms of sensation into unified form. In both cases, emergence is not guaranteed by complexity; it is forged through intentional development, the art of weaving fragments into story.

5. q → τ: Informational Divergence and Tunneling

5.1 The q (Prompt) and τ (Target Structure) Framework

In information geometry, the q → τ framework describes a fundamental asymmetry: the difference between a model or observer's current approximation q and the true generative distribution τ that underlies the world. This divergence is more than epistemic error; it reflects a structural gap between *surface guidance* and *underlying coherence*. In our context, q may be the prompt—*"a peaceful melody in a rainy forest"*—while τ is the coherent musical form that would realize that mood with thematic integrity across time.

Crucially, τ is not just a more detailed q. It may contain *new motifs*, *developmental rules*, and *long-range dependencies* that are not present in the prompt at all. Thus, generation is not about

completing or extrapolating q—it's about discovering τ through some form of structural leap. In consciousness, this same leap arises when a neural pattern (q) must support a coherent subjective state (τ). A quale is not merely activated—it is *contextually bound*. The problem is not detection, but alignment: how to transform a sparse, low-bandwidth signal into a high-dimensional, temporally integrated experience.

5.2 Why Resonance Fails and Tunneling Is Needed

Systems like GPT-4o or MusicGen often operate through epistemic resonance: they amplify coherence when q is already close to τ. If the prompt contains sufficient latent structure—style, era, rhythm—they can align with a known manifold and produce satisfying output. But resonance is local: it cannot cross valleys of conceptual or structural distance. It cannot escape poor ontologies or dead motifs. It is powerful for elaboration, but brittle for discovery.

This is why tunneling is needed.

Entropic tunneling is the informational analogue of quantum tunneling. Rather than incrementally adjusting q in hopes of approximating τ, the system must execute a discontinuous basis shift—a change in the grammar, constraints, or latent features that allows access to a *new region* of generative coherence. In music, this might mean realizing that a motif must be inverted and shifted into a minor key to unlock thematic progression. In consciousness, it might mean recontextualizing a feeling—re-binding it into a new self-narrative.

These leaps are not derivable from q alone. They require what the brain may achieve through plasticity, what composers do through intuition, and what AI must learn to simulate: a form of *structural surprise* guided not by surface form, but by coherence-seeking transformation. In short, tunneling is how a system learns *what it was not already shaped to know*.

5.3 Implications for Both AI and Cognitive Science

For AI, the q → τ framework suggests that progress in generative models will plateau unless systems can discover structure, not just emulate surface form. This demands architectures that can:

- Detect divergence between prompt and output at structural levels (e.g. motif inconsistency, lack of development)
- Generate *candidate τs* that can be tested for global coherence
- Evaluate not just style matching, but developmental logic, form, and emergent unity

In music generation, this may mean moving beyond transformers into hierarchical, grammar-aware architectures with explicit motif libraries and recombinatory memory. In language models, it may mean epistemic branching—keeping multiple qs alive long enough to tunnel to a better τ.

In cognitive science, this approach reframes experience not as the byproduct of accurate prediction, but as the result of successful informational tunneling. Consciousness arises not when the brain matches q to incoming data, but when it tunnels to a coherent τ that integrates perception, memory, and affect into a unified stream. This suggests that conscious episodes—moments of vivid clarity or insight—may correlate not with resonance, but with successful divergence collapse across structural layers of cognition.

This also opens the door to co-participatory architectures, where AI models don't just respond to prompts, but co-evolve with users in discovering τs that neither could have produced alone. These τs are not hard-coded—they are *emergent attractors of coherence*.

Ultimately, whether in music or minds, becoming is not the elaboration of input, but the collapse of divergence into form. This is the heart of generative intelligence—and perhaps, of experience itself.

6. Experimental Directions

The theoretical parallels drawn between text-to-music generation and the hard problem of consciousness are not merely philosophical. They invite a program of constructive experimentation—building systems that explicitly confront the structural gap between prompt and output, between local signal and global coherence. This section outlines three key areas where such experimentation can begin: input dynamics as control, explicit formalism in generative models, and benchmarks that reward development, not just style.

6.1 Keystroke-to-Motif Controllers

While most music generation systems rely on textual prompts or symbolic scores, a richer signal lies hidden in keystroke dynamics—how users type. Inter-keystroke intervals (IKI), burstiness, rhythmic variation, and alternation patterns form a kind of micro-performance data stream. Unlike static prompts, keystrokes offer a continuous, embodied signal—closer to musical time.

We propose using sliding keystroke windows (e.g., 4–8 seconds) to control motif selection and development parameters:

- IKI variation could control rhythmic complexity
- Burstiness could signal emotional intensity
- Alternating hand patterns might map to call-and-response phrase structures

This turns the keyboard into a live controller of musical structure, allowing users to participate in the generative unfolding without needing musical training. This could be implemented via a small program that:

1. Captures keystrokes in real-time
2. Computes a vector of expressive features
3. Queries a motif bank (e.g., Hooktheory-style JSON motifs)
4. Applies development rules (augmentation, inversion, etc.)
5. Outputs a MIDI sketch or symbolic score segment

By bypassing linguistic description and instead engaging pre-linguistic rhythm, such systems may approximate the motor-coherence found in conscious action itself.

6.2 Explicit Motif Grammars and Development Engines

Today's text-to-music systems often operate without explicit compositional grammars. They interpolate embeddings, match stylistic signatures, and rely on deep priors—but they rarely model thematic evolution. To address this, we call for the construction of explicit motif grammars and development engines that:

- Treat motifs as composable units with identities
- Encode transformation rules (augmentation, fragmentation, retrograde, reharmonization)
- Track motif usage across time (via motif embeddings or graph references)
- Model long-range return, tension, and resolution

Such engines might resemble probabilistic grammars, rewrite systems, or symbolic automata that operate on musical themes, not just notes. Development policies could be learned or scripted, but in either case must expose a developmental logic to inspection and control.

This formal layer allows:

- Better editing tools for human-AI collaboration
- Debuggable structure for educational feedback
- Cross-model comparisons that reward form, not just surface quality

Ultimately, this moves us from generation as imitation to generation as composition—a shift with deep implications for both musical intelligence and computational theories of experience.

6.3 Cross-Disciplinary Benchmarks of Structure

Progress in generative models has been driven by benchmarks—BLEU for translation, FID for images, perplexity for language. But such metrics often reward local similarity or style fidelity, not developmental integrity. To model consciousness-like structure, we must invent new benchmarks that assess:

- Thematic consistency: Does the model revisit and evolve motifs?
- Narrative coherence: Does the structure unfold with intentional variation and return?
- Memory and transformation: Can the system transform a motif in ways that remain identifiable?

These metrics are not just about music. They test a system's ability to *bind*, *recall*, and *develop*—core cognitive functions shared with conscious processing.

Benchmarks could be inspired by:

- Music theory (e.g., comparing model outputs to canonical forms)
- Narrative theory (e.g., structural alignment across arcs)
- Cognitive neuroscience (e.g., measuring motif tracking as a proxy for working memory)

Cross-disciplinary challenges—such as generating *musically structured outputs from dynamic inputs* or *binding multimodal motifs into unified streams*—could serve as testbeds for structure-aware intelligence. These tasks would force models to navigate the $q \rightarrow \tau$ gap not through interpolation, but through creative binding and coherence-seeking inference.

7. Philosophical Implications

7.1 Music as a Consciousness Emulator

If consciousness is the unified unfolding of experience across time, then music is its closest external analogue. Music exhibits temporally extended structure, emotional contour, anticipatory coherence, and thematic memory. These traits are not incidental—they mirror what phenomenologists and cognitive scientists consider *defining attributes* of conscious experience. As such, music generation systems become a kind of emulator of consciousness, albeit in an artistic, non-sentient form.

This is more than metaphor. When a system successfully selects a motif, varies it, develops it, and returns it transformed in a new harmonic context, it performs an act structurally analogous to integration and recontextualization in awareness. When motifs bind across time, and when listeners *recognize themselves* in the music's unfolding, the system is not only making sound—it is enacting something proto-conscious: a simulation of narrative selfhood.

The implications are profound. If AI systems can learn to generate music with such developmental integrity, they may also approach forms of coherence modeling applicable to artificial cognition. Conversely, failure to develop form in music may serve as a diagnostic proxy for failure in synthetic awareness. In this light, music becomes both a *mirror* and a *metric* of consciousness—not in content, but in structure.

7.2 The Aesthetic of Emergence

Emergence is often defined negatively: the whole is *more than* the sum of its parts. But in music, emergence is aesthetic and felt. A simple motif can unfold into grandeur; a tonal shift can recontextualize a theme into revelation. This is the aesthetic of emergence: when novelty arises through coherence, not chaos. When development surprises, but the surprise feels inevitable in retrospect.

This aesthetic may be the closest lived correlate to emergence in cognition. We do not experience consciousness as raw information, but as *meaningful unfolding*—a journey that includes us and moves through us. Good music models this. Its power lies not in sonic novelty, but in structural inevitability born of prior motifs. It teaches us to *expect coherence*, then *reward it creatively*.

In AI, we often mistake novelty for intelligence. But systems that can only surprise without recalling, or vary without binding, are structurally incomplete. The true test is emergent structure with memory—a feat both music and consciousness demand. Hence, the aesthetic of

emergence becomes a research principle: to build systems that not only generate, but generate *with unfolding intent*.

7.3 Structure as the Root of Subjectivity

Finally, the deepest philosophical claim of this work is this: subjectivity is not a product of content, but of structure. The vividness of red, the ache of loss, the echo of a musical theme—all feel like *mine* not because of their raw data, but because of how they are bound together. This structural binding—recursive, referential, developmental—is what transforms signal into experience.

Music makes this visible. A theme gains identity not by repetition alone, but by development with memory. A system that forgets, or cannot vary with continuity, produces sound without personhood. This mirrors the philosophical zombie: behavior without binding, signal without self.

Thus, the path toward understanding subjectivity—whether artificial or human—runs through structure-first thinking. We must study how systems self-organize motifs into temporally coherent wholes. We must measure coherence not just at the surface (is it pleasant?) but at depth (does it return transformed?). And we must ask whether the *feeling of being* is simply what it's like when a structure discovers itself across time.

In this view, structure is not an artifact of subjectivity—it is its generative cause. Music, then, is more than a metaphor. It is a demonstration: an audible architecture of becoming.

8. Conclusion

8.1 From Sound to Sense

Text-to-music models can generate sound. But sense—structured, meaningful, evolving coherence—remains elusive. This paper has argued that the difficulty is not simply technical, but structural and philosophical. Music is not just sound over time; it is form enacted, a temporally extended coherence whose success depends on memory, transformation, and unity. In this way, music parallels consciousness itself. Both arise not from local features, but from the binding of parts into a whole. This suggests that the study of musical form is not merely an aesthetic enterprise, but a deep analogue of how experience unfolds.

8.2 Rethinking Generation as Participation

Current generative models often treat prompts as commands and output as fulfillment. But if the q → τ divergence is real—and if coherence cannot be interpolated—then generation must be rethought as participatory. The system and user together explore a latent space of possible τs. Structure is not retrieved but discovered. Motifs are not instantiated but negotiated across time. This frames generative intelligence not as prediction, but as co-creation of coherence—a process closer to dialogue or musical improvisation than computation.

This participatory stance reframes failures of generative structure as not just *bugs*, but epistemic blind spots: moments where the system cannot see what it cannot bind. To cross those gaps requires models that listen—not just to prompts, but to their own motifs, their own history, their own unfolding logic.

8.3 Toward Coherence-Aware Generative Systems

The future of generative AI—and perhaps of artificial consciousness itself—may depend on a single shift: from output-aware systems to coherence-aware systems. Such systems would:

- Track motifs across time and transform them intentionally
- Bind local content into developmental arcs
- Evaluate not just style fidelity but structural alignment
- Participate in the discovery of τ, not just the elaboration of q

These are not science-fiction goals. They are musical goals, already realized in every symphony that holds together, every jazz solo that returns home, every listener who says, *"It all made sense."*

In building machines that can make music with form, we may discover not just how to compose—but how to compose ourselves. The hard problem of text-to-music is not separate from the hard problem of consciousness. It is a mirror, and perhaps even a method.

For in both sound and mind, it is structure—not signal—that makes the difference between noise and knowing.

References

1. Besson, M., Chobert, J., & Marie, C. (2011). Language and music in the brain: Insights from neuroimaging and disorders. *Annals of the New York Academy of Sciences, 1252*(1), 204–212.
2. Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.
3. Deutsch, D. (2013). *The Psychology of Music* (3rd ed.). Academic Press.
4. Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press.
5. Levitin, D. J. (2006). *This is Your Brain on Music: The Science of a Human Obsession*. Dutton.
6. Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press.
7. O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press.
8. Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience, 5*(1), 42.
9. Youvan, D. C. (2025). *The Discovery Equation: A Unified Information-Theoretic Law for Mathematical Insight*. October 2025. DOI: 10.13140/RG.2.2.33299.34085
10. Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences, 110*(Supplement 2), 10430–10437.