

Calamus Sound Engine: Foundations v2.1

Note: this document contains early exploration of concepts that may have changed later on in other specifications documents

The Architectural Fork

Before implementation, a fundamental choice: how to generate sound?

Additive approach:

- Build from sine components
- Precise control over each harmonic
- Richness must be constructed — every bit of complexity is intentional
- Clean, surgical, but life must be explicitly programmed

Dough approach:

- Start with something already rich
- Sculpt by subtraction and transformation
- Complexity is inherent — you tame it rather than build it
- Less precise, more organic

Decision: Dough.

The reasoning: experience shows that explicitly programming life in the digital domain rarely succeeds. The better path is working *with* a medium that has inherent tendencies — like painting with pigments that want to bleed, on paper that wants to absorb. You guide more than dictate. The sound emerges.

The DNA Metaphor

Living material needs DNA. Not a specification of every cell, but instructions for *how to build* and *how to respond*. The organism emerges from DNA interacting with environment over time.

Sound DNA encodes two gene types:

Identity Genes (what makes it "this instrument")

- Harmonic recipe — relationships and tendencies, not fixed amplitudes
- Formant regions — resonances that persist across pitch
- Noise character — the breath between harmonics
- Coupling rules — when one parameter moves, what follows?

Behavioral Genes (how it responds to gesture)

- Mass/inertia profiles for each parameter
- How energy translates to spectral richness
- Attack characteristics — how it begins
- Release behaviors — how it lets go

Family and Individual

- **Family genes** — shared by all four voices, the core identity
- **Individual expression** — register placement, formant tuning, behavioral variations (bass more massive, soprano more agile)

The gesture becomes the environment. DNA meets gesture, and sound is *expressed* — in the biological sense.

First Principles: What Is Sound?

Setting aside musical instruments. Pure physics.

What is sound? A vibration propagating through a medium.

What produces vibration? A material disturbed from equilibrium, with restoring forces pulling it back, overshooting, oscillating.

What determines vibration character?

The material's properties:

- **Mass** — resistance to acceleration
- **Elasticity** — strength of restoring force
- **Geometry** — shape, boundaries, constraints
- **Damping** — energy loss over time

These determine the material's **modes** — preferred patterns of vibration.

What about excitation?

Excitation is energy entering the system. It doesn't create modes — it *activates* them. The same string bowed gently or bowed hard near the bridge produces different balances of the same underlying modes.

The Deep Structure

Material → Modes → Identity

Excitation → Mode activation → Expression

The DNA *is* the modal structure. What vibration patterns are possible, how they relate, how they decay, how they couple.

We are not modeling a violin. We are **inventing a material** with desired modal properties.

What Is Most Fundamental?

Initial assumption: harmonic signature is most fundamental.

But consider: a violin bowed near the bridge versus over the fingerboard produces dramatically different harmonic signatures. Yet both are unmistakably violin. What persists?

The body.

The resonant cavity. The formants. How the physical structure colors *whatever passes through it*.

The violin string's vibration varies wildly with bow position, pressure, speed. But all passes through the same wooden body with fixed resonances. The body is constant. Excitation is variable.

This suggests: identity lives primarily in the **filter** (formant fingerprint) rather than in the harmonic signature of excitation.

Harmonic signature may be *emergent* — result of source meeting filter — rather than fundamental.

Implication: The DNA's deepest gene is: *how does this body resonate?*

Lab discovery: In practice, harmonics + rolloff appear to define core identity even before formants. The formants refine and add vowel character, but the raw dough already has family coherence.

The Desired Material

Aesthetic goals (the tailor shop moment):

- Rich in timbral variations
- Agile in upper registers
- Character somewhere between violin, cello, saxophone, human voice
- Able to produce vowels

Translated to fundamental properties:

Aesthetic Goal	Fundamental Property
Rich in timbral variations	Many modes, densely packed, complex relationships
Agile in upper registers	Lower mass/inertia for higher modes, quick response
Violin/cello/sax/voice character	Pronounced formants — shaped resonance, not flat response
Vowel capability	Moveable formants — resonances that shift under gesture control

The Dough Recipe — Proven in Lab

Base Ingredients

Modal richness (Harmonics) Dense spectrum. Many partials (up to 64 tested). The number of harmonics alone creates progression from simple (OH vowel at 4) to complex.

Spectral slope (Rolloff) The "magic ingredient." Controls how quickly upper harmonics decay. Low rolloff = bright/buzzy. High rolloff = dark/hollow. This parameter makes harmonic count meaningful across the entire range — without it, sound becomes static "lawn mower" above ~15 harmonics.

Formant body (F1 + F2) Two resonant bandpass filters that impose throat/body character. F1 (200-1000 Hz) roughly corresponds to jaw openness. F2 (500-3000 Hz) corresponds to tongue position. Together they create vowel space.

Breath / Turbulence (Voice-Noise Blend) Not simple "noise added on top" — but a continuum between harmonic content and shaped turbulence. Both pass through the same formant body. The breath component is what makes sound human.

Blend	Character
0%	Pure electronic, vowels
5-15%	Human voice quality — lungs, throat, body
30-50%	Breathy, intimate (Ben Webster territory)
70-90%	Whispered, voiced fricatives (Z, V)
100%	Pure fricative (S, SH, F)

Noise Color Shapes the turbulence spectrum before it hits formants:

- Low (100 Hz cutoff) = dark, breathy, rumble
- Mid (2000 Hz) = SH territory
- High (8000 Hz) = S territory, hissy

Micro-life (Drift) Subtle detuning of each harmonic ($\pm 0.5\%$). Creates slow beating patterns between partials. At low values, produces vibrato-like character without explicit vibrato. Adds organic quality.

Dynamic Properties (Behavioral Genes)

Mass/Inertia Parameters resist instant change. Springs pull toward targets, velocity accumulates, damping prevents endless oscillation. Higher mass = slower response, more overshoot, heavier feel.

Critical discovery: Mass isn't just expressive — it protects the system. Filter coefficient discontinuities cause audio glitches; mass smooths transitions automatically. Physics provides stability.

Coupling Parameters are connected, not independent. Implemented first between F1 and F2:

- **Positive coupling:** Parallel motion. When F1 moves up, F2 is attracted upward. Vowels stay coherent.
- **Negative coupling:** Contrary motion. When F1 moves up, F2 is attracted downward (mirror). Vowels become exaggerated, vocabulary expands.
- **Zero:** Independent movement.

Coupling model evolved through three iterations:

1. Distance-based forces → runaway instability
2. Target distance maintenance → stuck at extremes
3. Target blending with mirror → stable, musical, dynamic ✓

Pitch Glide Fundamental frequency has physics too — mass/spring/damping. Moving the pitch target causes the actual pitch to glide rather than jump. Higher mass = slower, more expressive portamento. The same physics that works for formants works for pitch.

DNA Sequences: Harmonic Distribution

Major Discovery: The harmonic distribution is the genetic code itself.

Different distributions create fundamentally different instrument families, all responding to the same physics:

Sequence	Pattern	Character
All	1,2,3,4,5...	Full harmonic series — rich, complex

Sequence	Pattern	Character
Odd Only	1,3,5,7...	Clarinet-like, closed pipe
Even Only	2,4,6,8...	Hollow, ethereal
Odd Dominant	Odds 100%, evens 30%	Fuller hybrid
Fundamental+	Strong 1, weak overtones	Pure, simple, near-sine
Octaves Only	1,2,4,8,16,32...	Organ-like, hollow power

Purity control: Blend between pure distribution and "All harmonics" — species can hybridize.

Implication: Calamus can have infinite instrument families, each with four voices, each responding to the same gesture vocabulary. The DNA sequence determines species; the physics determines expression.

Physics-Based Attack: The Breakthrough

The Wrong Path: Noise

Initial attempt: add noise burst for attack transient. Result: "small dog barking." Noise sits on top of the sound — doesn't integrate. Even filtered through formants, it feels "added."

The Right Path: Energy Injection

Attack is not noise added to the signal. Attack is **energy injected into the physics**.

What happens on attack trigger:

1. **Formant kick** — F1 and F2 receive velocity impulse. They lurch from their rest position, then spring/damping brings them back. The *movement* of formants = attack character.
2. **Brightness bloom** — Rolloff temporarily decreases (more upper harmonics), then decays back to normal. Brighter attack → mellower sustain.
3. **Amplitude swell** — Brief boost, then settles.

No noise. The attack is the system *responding* to energy, not a sample pasted on.

Formant Velocity: Consonants Emerge

The "Ba Ba" Milestone:

By adding direct control over F1 and F2 velocity at attack:

- **F1 Velocity: -500** (formant sweeping down then up)
- **F2 Velocity: +270** (contrary motion)
- **Coupling: 56%** (connected but not locked)
- **Mass: 15%** (quick, responsive)
- **Rolloff: 91%** (dark, vocal)

Result: The system said "**Ba ba.**"

Real consonants from physics. The "B" is the formant trajectory — starting displaced, sweeping into vowel position. No samples. No synthesis tricks. Just physics.

Implication: Consonant space exists within formant velocity parameters. "B", "M", "D", "G" — they're all trajectories waiting to be found.

Gate and Release: Birth and Death

The infant learned to begin and end sounds.

Gate mechanics:

- **Gate On (key down):** Triggers attack (formant velocity kick + brightness bloom), amplitude rises
- **Gate Off (key up):** Amplitude fades out gently

Musical implications:

- Hold for sustained notes
- Tap for staccato
- The release is as expressive as the attack

Future: Envelope per parameter Currently amplitude has attack/release. But expressively:

- Voice/Noise could swell (breath increases as note dies — Ben Webster's breathy release)

- Brightness could evolve (attack bloom already exists)
- Formants could shift (vowel evolution within a note)

Envelopes = the path to phrasing. Parked for fabrication phase.

The Breath Continuum: What Makes Sound Human

Key Discovery: A touch of turbulence (5-15% Voice/Noise blend) transforms electronic vowels into something that sounds like it has lungs, a throat, a body.

The breath is what makes us human.

Character	Where It Lives
Pure electronic	0% blend
Human voice	5-15% blend
Intimate/breathy	20-40% blend
Whispered	50-70% blend
Fricative	80-100% blend

The space in between: Calamus doesn't have to be a human voice OR a traditional instrument. It can move along the continuum, finding new territories that have qualities of both.

This connects to the original vision: "*Character somewhere between violin, cello, saxophone, human voice.*" The breath component is the connective tissue.

The Embryo Becomes Infant: Current State

The metaphor evolved:

- **Dough** → raw material with properties
- **Embryo** → living material that develops and responds
- **Infant** → speaks, breathes, has voice, personality emerging

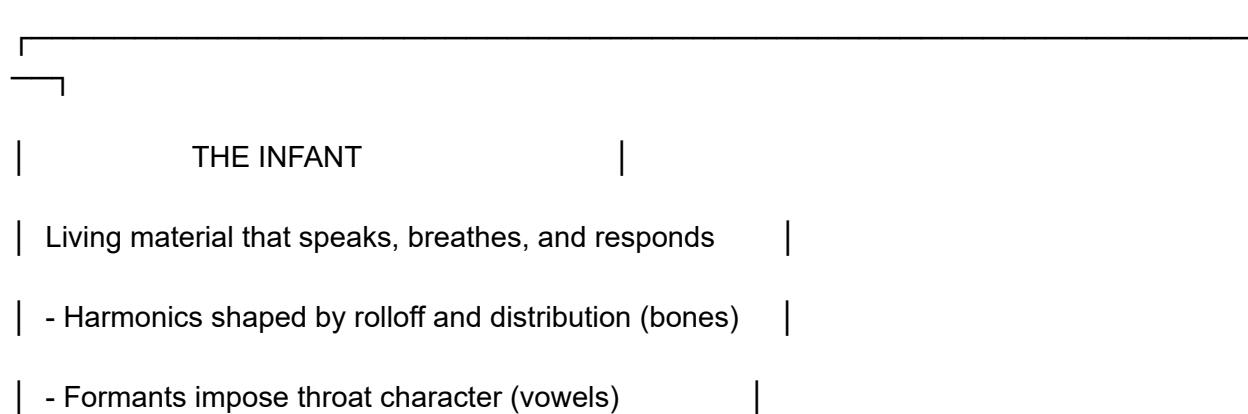
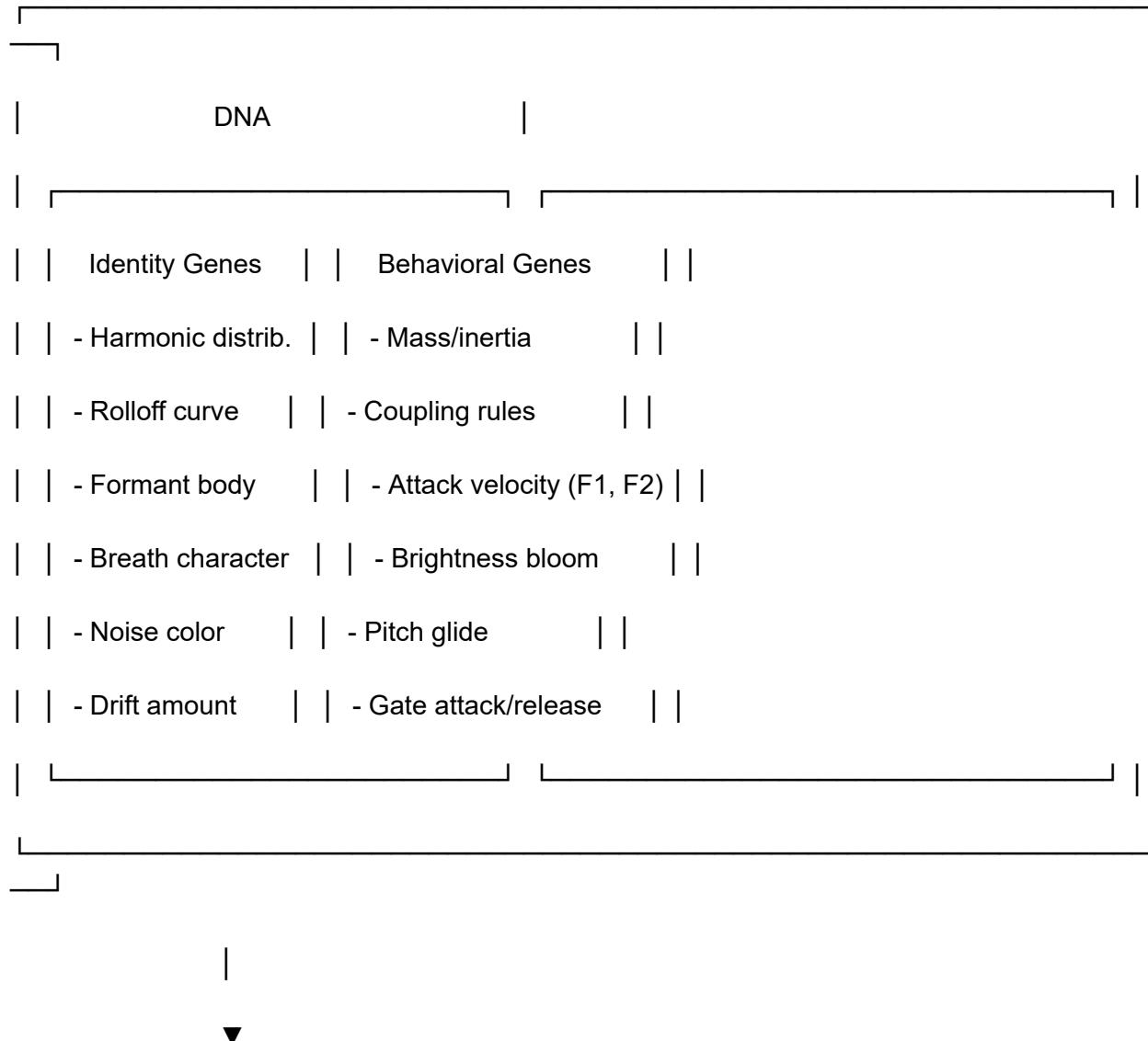
Anatomy Complete

Structure	Implementation	Status
Bones	Harmonics + rolloff	✓ Core identity
DNA	Harmonic distribution	✓ Infinite species
Lungs	Voice/Noise blend	✓ Breath continuum
Breath color	Noise filter	✓ Fricative spectrum
Heartbeat	Drift/detuning	✓ Subtle life
Throat	F1 + F2 formants	✓ Vowels
Weight	Mass/inertia	✓ Physics + stability
Tendons	F1-F2 coupling	✓ Parallel/contrary motion
Voice	Brightness bloom	✓ Attack dynamics
Speech	Formant velocity	✓ Consonants — "Ba ba"
Birth/Death	Gate + release	✓ Notes begin and end
Glide	Pitch physics	✓ Portamento

Emergent Properties Discovered

1. **Family coherence without explicit design** — Raw harmonics + rolloff create "same instrument" quality
 2. **Vowel-like sounds emerge** — Low harmonic counts produce OH→AH progression
 3. **Rich syllable vocabulary** — Coupling creates range of possible sounds
 4. **Physics enables stability** — Mass smooths discontinuities that cause glitches
 5. **Infinite species** — DNA sequences create distinct instrument families
 6. **Consonants from physics** — Formant velocity trajectories = speech-like articulation
 7. **"Ba ba"** — The system learned to speak
 8. **Breath = humanity** — Small turbulence component transforms electronic into organic
 9. **The space in between** — Neither human voice nor traditional instrument, but new territory
-

Architecture Summary



- | - Formant velocity creates consonants (speech) |
- | - Parameters have mass and coupling (weight, tendons) |
- | - Brightness blooms on attack (voice) |
- | - Pitch glides with physics (portamento) |
- | - Breath/turbulence adds humanity (lungs) |
- | - Gate controls birth and death (life cycle) |



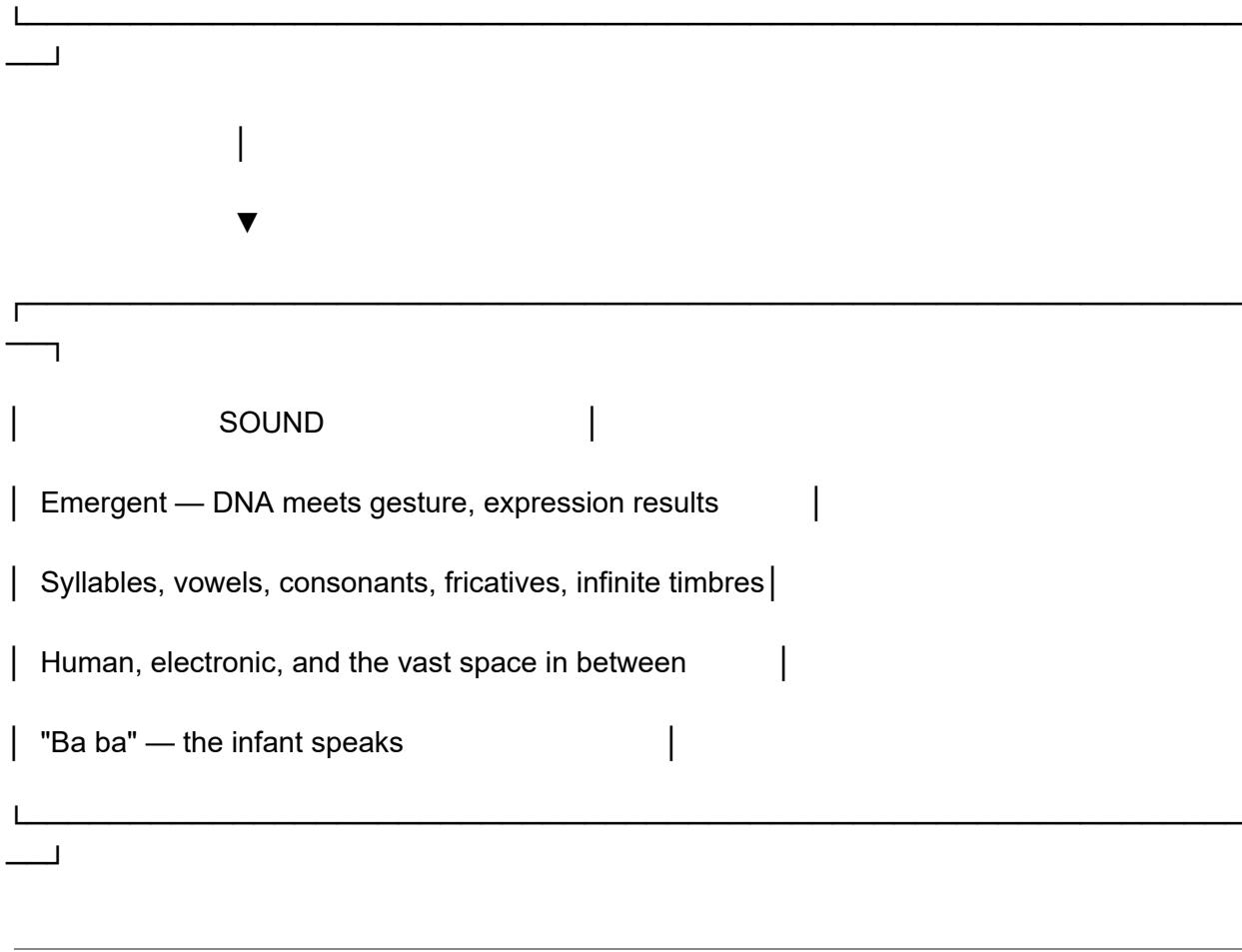
|

▼



GESTURE

- | Energy entering the system |
- | (Currently: sliders + spacebar. Future: Wacom pen) |
- | - Formant position (vowel shaping) |
- | - Formant velocity (consonant shaping) |
- | - Voice/Noise blend (breath amount) |
- | - Noise color (fricative character) |
- | - Attack trigger / Gate (energy injection) |
- | - Pitch control (with glide) |
- | - Brightness/rolloff modulation |



Lab Infrastructure

Platform: C# / WinForms / NAudio

Architecture: Real-time synthesis via ISampleProvider

Key classes:

- **DoughProvider** — Sample-by-sample synthesis with physics
- **ResonantFilter** — Biquad bandpass for formants
- **NoiseColorFilter** — High-pass filter for turbulence shaping

Parameters controllable in real-time:

Parameter	Range	Function
Harmonics	1-64	Modal richness

Parameter	Range	Function
Voice/Noise	0-100%	Breath continuum
Noise Color	0-100%	Fricative spectrum (dark→bright)
Drift	0-100%	Micro-detuning for life
Frequency	50-500 Hz	Fundamental pitch (with glide)
Rolloff	10-100%	Brightness / spectral slope
F1	200-1000 Hz	First formant (jaw)
F2	500-3000 Hz	Second formant (tongue)
Mass	0-100%	Physics weight
Coupling	-100 to +100%	F1-F2 relationship
DNA Sequence	Dropdown	Harmonic distribution
Purity	0-100%	Species hybridization
F1 Velocity	-500 to +500	Consonant attack shape
F2 Velocity	-500 to +500	Consonant attack shape
Gate	Spacebar hold	Note on/off

Open Frontiers (For Fabrication)

Immediate priorities:

1. **Parameter envelopes** — Attack/sustain/release per parameter, not just amplitude. The path to phrasing (Ben Webster's breathy release).
2. **Four voices** — Same DNA, different register placement and formant tuning. Soprano, Alto, Tenor, Bass.

3. **Gesture integration** — Replace sliders with Wacom pen input. Six dimensions of continuous control.
4. **Preset system** — Snapshots of full state as departure points. Not "sounds" but "starting places to deform from."

Later:

5. **Custom DNA sequences** — User-defined harmonic distributions beyond the presets.
 6. **Tensor architecture** — Multiple behavior layers with individual characteristics.
 7. **Consonant mapping** — Systematic exploration of velocity space for full phoneme vocabulary.
 8. **Test melody** — Looped phrase exercising: long notes, short notes, legato, staccato, chords.
-

Key Principles Confirmed

1. **Dough > Additive** — Working with rich material beats assembling from parts
 2. **Physics is structural** — Behavioral genes provide both expression AND stability
 3. **Identity emerges** — Family coherence appears without explicit design
 4. **Coupling expands vocabulary** — Connected parameters create more possibilities than independent ones
 5. **DNA sequences = species** — Harmonic distribution is the genetic code
 6. **Attack = energy injection** — No noise, just physics responding
 7. **Formant velocity = consonants** — Trajectory through vowel space creates speech
 8. **Breath = humanity** — Turbulence through formants makes sound alive
 9. **The space in between** — Not voice, not instrument, but new territory
 10. **The ear is the judge** — Lab experiments reveal what specs cannot predict
-

The Painting Insight

From Nimus, regarding Rebelle painting software:

"You just put blobs of color that you kinda control but not really and you pile them up until the picture emerges. And it does."

This remains the target experience. The sound engine feels like a responsive medium with inherent tendencies. Guide more than dictate. The sound emerges.

Milestone Log

Session 1:

- Established dough vs. additive decision
- Built harmonics + rolloff → "lawn mower" → discovered rolloff as magic ingredient
- Added formants → "strange language" → throat emerged
- Added mass/inertia → physics + stability discovery
- Added coupling → three iterations → parallel/contrary motion works

Session 2:

- Added DNA sequences (harmonic distribution) → infinite species
 - Attempted noise-based transient → "chihuahua barking" → wrong path
 - Reframed as physics-based attack → "boueh, piehhh, aweuuuhh" → right path
 - Added formant velocity control → "**Ba ba**" → the infant speaks
 - Added gate/release → notes begin and end
 - Added pitch glide → portamento through physics
 - Added Voice/Noise blend + Noise Color → fricatives and breath continuum
 - Discovery: breath makes sound human, 5-15% transforms electronic to organic
-

Ready for Fabrication

The lab phase is complete. We have:

- ✓ A working sound engine with all core components
- ✓ Physics-based parameter behavior
- ✓ Infinite instrument families via DNA sequences
- ✓ Consonant articulation from formant velocity
- ✓ The breath continuum from pure tone to fricative
- ✓ Gate/release for note lifecycle
- ✓ Pitch glide for expressive portamento
- ✓ Documented architecture and principles

Next phase: Build the real application — the Input Engine that captures Wacom gesture and drives this sound engine. The pen becomes the breath.

The Name

Calamus — Latin for both "reed" (wind instrument) and "pen" (writing tool). The two worlds unified.

The sound engine is the reed: where gesture becomes voice, where physics becomes music.

Status: Lab complete. The infant speaks, breathes, and is ready to leave the incubator.

Document version: 2.0

Lab sessions: December 2025

"Ba ba" milestone: December 2025

Breath continuum discovery: December 2025

Ready for fabrication 